The MERRA-2 Aerosol Reanalysis, 1980 – onward, Part I: System Description and Data Assimilation Evaluation

C. A. Randles [*][a][†] A. M. da Silva [a], V. Buchard[a,b], P. R. Colarco[c], A. Darmenov[a], R. Govindaraju[a,d], A. Smirnov[d,e], B. Holben[e], R. Ferrare[f], J. Hair[f], Y. Shinozuka[g,h] and C. J. Flynn[i]

[a]*Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA*
[b]*GESTAR/Universities Space Research Association, Columbia, Maryland, USA*
[c]*Atmospheric Chemistry and Dynamics Lab, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA*
[d]*Science Systems and Applications, Inc., Lanham, MD, USA*
[e]*NASA Biospheric Sciences Laboratory, Greenbelt, Maryland, USA*
[f]*NASA Langley Research Center, Hampton, VA, USA*
[g]*Bay Area Environmental Research Institute, Petaluma, California, USA*
[h]*NASA Ames Research Center Cooperative for Research in Earth Science and Technology, Moffett Field, California, USA*
[i]*Pacific Northwest National Laboratory, Richland, Washington, USA*

ABSTRACT

The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2) updates NASA's previous satellite era (1980 – onward) reanalysis system to include additional observations and improvements to the Goddard Earth Observing System, Version 5 (GEOS-5) Earth system model. As a major step towards a full Integrated Earth Systems Analysis (IESA), in addition to meteorological observations, MERRA-2 now includes assimilation of aerosol optical depth (AOD) from various ground- and space-based remote sensing platforms. Here, in the first of a pair of studies, we document the MERRA-2 aerosol assimilation, including a description of the prognostic model (GEOS-5 coupled to the GOCART aerosol module), aerosol emissions, and the quality control of ingested observations. We provide initial validation and evaluation of the analyzed AOD fields using independent observations from ground, aircraft, and shipborne instruments. We demonstrate the positive impact of the AOD assimilation on simulated aerosols by comparing MERRA-2 aerosol fields to an identical control simulation that does not include AOD assimilation. Having shown the AOD evaluation, we take a first look at aerosol-climate interactions by examining the shortwave, clear-sky aerosol direct radiative effect. In our companion paper, we evaluate and validate available MERRA-2 aerosol properties not directly impacted by the AOD assimilation (e.g. aerosol vertical distribution and absorption). Importantly, while highlighting the skill of the MERRA-2 aerosol assimilation products, both studies point out caveats that must be considered when using this new reanalysis product for future studies of aerosols and their interactions with weather and climate.

1. Introduction

The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2) is the new modern satellite era (1980 – onward) atmospheric reanalysis from the NASA Global Modeling and Assimilation Office (GMAO, Gelaro et al. 2017). Following the success of the original MERRA reanalysis (Rienecker et al. 2008, 2011), the MERRA-2 system incorporates new observations not available for MERRA and reduces spurious trends and jumps related to changes in the meteorological observing system (McCarty et al. 2016). Numerous improvements have been made to the MERRA-2 Goddard

Earth Observing System, Version 5 (GEOS-5) modeling system (Molod et al. 2012, 2015), including improvements in the representation of the hydrologic cycle (Takacs et al. 2015; Reichle and Liu 2014), the stratosphere, ozone, and cryospheric processes (Bosilovich et al. 2016).

In a significant step towards an Integrated Earth Systems Analysis (IESA), MERRA-2 for the first time includes analyzed aerosol fields that are radiatively coupled to the atmosphere. To our knowledge, this is the first multi-decadal reanalysis within which meteorological and aerosol observations are jointly assimilated into a global assimilation system, though other operational forecasting centers are actively developing similar capabilities (e.g. Benedetti et al. 2009; Sekiyama et al. 2010; Lynch et al. 2016). Previously, the GMAO had performed an off-line aerosol reanalysis (MERRAero) in which bias-

─────────

[*]*Corresponding author address:* C. A. Randles, 1545 US-22, Clinton, NJ, 08801

E-mail: cynema@alum.mit.edu

[†]now at ExxonMobil Research and Engineering

corrected MODerate resolution Imaging Spectroradiometer (MODIS) aerosol optical depth (AOD) from Terra and Aqua was assimilated into an earlier version of the NASA GEOS-5 model driven by meteorology from MERRA-1 (Buchard et al. 2015, 2016). In addition to assimilating bias-corrected Collection 5 MODIS AOD as in MERRAero, MERRA-2 now also includes assimilation of bias-corrected AOD from the Advanced Very High Resolution Radiometer (AVHRR) instruments (Heidinger et al. 2014), AOD retrievals from the Multi-angle Imaging SpectroRadiometer (MISR) over bright surfaces (Kahn et al. 2005), and ground-based AErosol RObotic NETwork (AERONET) direct measurements of AOD (Level 2, Holben et al. 1998).

Despite their rapid increase in complexity, aerosol models remain uncertain due to poorly constrained emissions and physical process parameterizations such as hygroscopic growth, mixing, and aerosol-cloud interactions (Textor et al. 2006; Kinne et al. 2006; Benedetti et al. 2009; Schutgens et al. 2010). Similarly, many aerosol observations such as those from remote sensing platforms, both satellite- and ground-based, suffer from limited coverage (e.g. due to their orbit and/or cloud contamination), contextual biases such as "clear-sky" bias, and biases due to assumptions made in retrieval algorithms (Zhang and Reid 2009; Shi et al. 2011; Colarco et al. 2014). Reanalyses attempt to take advantage of the best features of both models and observations to produce four-dimensional, gridded output that optimally combines the continuity of a model with real-world observations that may be sparse and/or irregularly spaced both spatially and temporally (Rienecker et al. 2011; Schutgens et al. 2010). The analyzed aerosol fields from reanalyses such as MERRA-2 have numerous applications (see Bocquet et al. (2015) and citations therein). Briefly, these fields can serve as initial conditions for regional modeling and air quality forecasting (Zhang et al. 2012; Giordano et al. 2015; Buchard et al. 2016), as a tool to investigate aerosol-climate or aerosol-weather interactions (Bellouin et al. 2013; Reale et al. 2014), for use as *a priori* profiles used in satellite retrievals of other atmospheric constituents (Kessner et al. 2013; Inness et al. 2013), and for optimal network/satellite sensor design in the context of Observing System Simulation Experiments (OSSEs; Bocquet et al. 2015).

Because standard minimum variance data assimilation algorithms are designed to minimize random errors under the assumption of no biases, systematic errors in model background and observations must be carefully accounted for, or, to the extent possible, removed prior to the assimilation process (Dee and da Silva 1999; Zhang and Reid 2009; Benedetti et al. 2009). On the observational side, this can be accomplished by imposing strict quality control on the observing system used by the reanalysis (Zhang and Reid 2006). Additionally, it is essential to have a well-performing and well-tuned prognostic model to minimize

the corrections imposed by the assimilation since large corrections can cause spurious trends and jumps in the analyzed fields when the observing system changes over time (Lynch et al. 2016).

Several important caveats must be understood when using and evaluating the MERRA-2 aerosol reanalysis products. First, the relative paucity of (non-assimilated) aerosol observations make independent validation of the analyzed AOD a challenge. Second, despite best efforts at harmonizing the observing system through quality control, differences in data coverage can and do impact the analyzed AOD, particularly between the pre- and post-NASA Earth Observing System (EOS) periods (1980 – 1999 and 2000 – onwards, respectively). Finally, non-analyzed aerosol properties (e.g. vertical distribution, aerosol speciation, absorption) are not fully-constrained by the assimilation and draw strongly to the assimilating model in most cases. Nevertheless, despite some deficiencies, previous studies (e.g. Buchard et al. 2015, 2016), the current study, and a companion evaluation paper (Buchard et al. 2017) demonstrate that the aerosol assimilation system does indeed show considerable skill in simulating numerous observable aerosol properties.

The purpose of this study is to describe the MERRA-2 aerosol data assimilation system, provide initial validation of the analyzed AOD fields, and suggest applications of the aerosol products while highlighting both the model's skill and deficiencies. In Section 2, we provide details about the GEOS-5/GOCART model and aerosol emissions. Section 3 discusses the aerosol assimilation system and the AOD observing system. We then evaluate the performance of the analyzed AOD fields with respect to the observing system and demonstrate the stability of the assimilation system (Section 4a). The impact of the AOD assimilation is shown by comparisons to a control simulation driven by MERRA-2 meteorology with identical aerosol emissions, but without AOD assimilation (Section 4b). In Section 4c, we present a validation of MERRA-2 AOD with available, non-assimilated observations. Finally, we examine the MERRA-2 clear-sky, short-wave aerosol direct radiative effect (DRE, Section 4d) and compare to results from other AOD reanalyses. A summary is presented in Section 5.

In Part II of this study, Buchard et al. (2017) present an evaluation of the MERRA-2 system with respect to non-assimilated aerosol properties (e.g. absorption, vertical profile, $PM_{2.5}$) and case studies during major aerosol events (e.g. the 1991 Pinatubo eruption). In both this work and Buchard et al. (2017), we demonstrate important aspects of the aerosol products that users must consider. We note that the full MERRA-2 reanalysis dataset including aerosol fields is currently publicly available online through the Goddard Earth Sciences Data and Information Services Center (GES DISC; `http://disc.`

sci.gsfc.nasa.gov/mdisc/) with technical documentation (Randles et al. 2016) and file specifications available at http://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/docs/. Individual MERRA-2 data collections are accessible via their own digital object identifier (doi) codes. For this study and Buchard et al. (2017) we use aerosol (Global Modeling and Assimilation Office 2015a,b,c), meteorological (Global Modeling and Assimilation Office 2015d,e), and radiation (Global Modeling and Assimilation Office 2015f) collections.

2. The MERRA-2 Modeling System

An overview of the MERRA-2 modeling system is found in Gelaro et al. (2017). Briefly, MERRA-2 is produced using the GEOS-5 atmospheric model and data assimilation system version 5.12.4 (Rienecker et al. 2008; Molod et al. 2015) and the Three-dimensional Variational Data Analysis (3DVAR) Gridpoint Statistical Interpolation (GSI) meteorological analysis scheme (Wu et al. 2002; Kleist et al. 2009). The GSI uses an incremental analysis update procedure every 6 hours (Bloom et al. 1996). The discretization of the dynamical core (Lin 2004) is computed on a cubed sphere grid that mitigates grid-spacing singularities (Putman and Lin 2007). The GEOS-5 model resolution on the native cubed-sphere grid is roughly 50 km with 72 hybrid-eta layers from the surface to 0.01 hPa, while most products are saved on a standard $0.5° \times 0.625°$ latitude by longitude grid. As noted previously, the MERRA-2 meteorological observing system includes numerous additions that are detailed in McCarty et al. (2016), and Bosilovich et al. (2016) presents the validation of the MERRA-2 meteorological, radiation, ozone, and cryospheric fields. The remainder of this section focuses on aspects of the MERRA-2 modeling system relevant for the aerosol reanalysis.

a. GOCART Aerosol Module

Aerosols in MERRA-2 are simulated with a radiatively coupled version of the Goddard Chemistry, Aerosol, Radiation, and Transport model (GOCART, Chin et al. 2002; Colarco et al. 2010). GOCART treats the sources, sinks, and chemistry of 15 externally mixed aerosol mass mixing ratio tracers: dust (5 non-interacting size bins), sea salt (5 non-interacting size bins), hydrophobic and hydrophilic black and organic carbon (BC and OC, respectively; 4 tracers), and sulfate (SO_4). Both dust and sea salt have wind-speed dependent emissions. Primary sulfate and carbonaceous aerosol species have emissions principally from fossil fuel combustion, biomass burning, and biofuel consumption, with additional biogenic sources of organic carbon. Secondary sources of sulfate include chemical oxidation of sulfur dioxide gas (SO_2) and di-methyl sulfide

(DMS), and we include a database of volcanic SO_2 emissions and injection heights. Note that we use a monthly-mean climatology of oxidant fields in GOCART. Loss processes for all aerosols include dry deposition (including gravitational settling), large-scale wet removal, and convective scavenging. We note that precipitation-induced aerosol deposition does not depend on model-generated precipitation fields. Rather, we use the MERRA-2 corrected precipitation product of Reichle et al. (2017) that better represents diurnal precipitation changes compared to observations. Aerosol hygroscopic growth depends on simulated relative humidity and is considered in computations of particle fall velocity, deposition velocity, and optical parameters. Numerous studies have demonstrated the skill of the GOCART aerosol module in simulating AOD and other observable aerosol properties (e.g. Colarco et al. 2010; Nowottnick et al. 2010, 2011; Bian et al. 2013).

b. Emissions

Figure 1 shows the EOS-period climatological distribution of aerosol emissions in MERRA-2 from GOCART. Dust emissions (Fig. 1a) use a map of potential dust source locations based on the observed correlation of dust emitting regions with large-scale topographic depressions (Ginoux et al. 2001, updated to $0.25°$ resolution). Emissions of both dust and sea salt (Fig. 1b) are wind-driven for each size bin, parameterized following Marticorena and Bergametti (1995) and Gong (2003), respectively. Sea salt emission strength is modulated with a sea surface temperature (SST) derived correction similar to the work of Jaeglé et al. (2011). The wind dependent term in the Gong (2003) sea salt emission parameterization is updated to depend on friction velocity rather than 10-m velocity (i.e. $u_*^{2.41}$, where u_* is the friction velocity). We note that there are errant emissions of sea salt over the Great Lakes, Caspian and Black Seas due to an issue in the lake-masking algorithm during the simulation. While there is minimal impact of lake emissions on the assimilated AOD fields, these inland sea salt emissions have consequences for other aerosol products such as speciated $PM_{2.5}$ in some locations (Buchard et al. 2017; Randles et al. 2016).

Sulfate and carbonaceous aerosol emissions (Fig. 1c-f) derive from both natural and anthropogenic sources. Table 1 outlines the emission inventories with additional details given in Randles et al. (2016). With the exception of volcanic SO_2 emissions, when a given emissions inventory ends the final emission year is persisted in the model. Natural emissions of SO_2 from volcanoes derive from the AeroCom Phase II project (Diehl et al. 2012, http://aerocom.met.no/) and cover eruptive and degassing volcanoes on all days from January 1, 1979 to December 31, 2010. Only a repeating annual cycle of degassing volcanoes is included in MERRA-2 after 2010. Eruptive volcanoes emit in the upper third of the column

defined by the volcano elevation and the plume-top height; degassing volcanoes emit at the volcano elevation. With the exception of aircraft and energy-sector sulfur dioxide, anthropogenic aerosol sources emit into the lowest model-layer. Energy-sector emissions of SO_2 (EDGARv4.2, European Comission 2011) are emitted between 100 and 500 m above the surface as in Buchard et al. (2014).

Owing to the intense vertical mixing associated with fires, we distribute biomass burning emissions uniformly throughout the planetary boundary layer (PBL) in the grid box where the fire emission occurs. A diurnal cycle, which is more prominent in the tropics and gradually weakens in the higher latitude extratropical temperate zones, is imposed online for biomass burning emissions. Importantly, biomass burning emissions of carbonaceous and sulfate aerosols in MERRA-2 derive from a variety of inventories over the course of the reanalysis. Figure 2 shows global and regional timeseries of the carbonaceous aerosol biomass burning emissions.

From 2010, daily emissions of biomass burning OC, BC, and SO_2 derive from the Quick Fire Emissions Dataset (QFED) version 2.4-r6 (Darmenov and da Silva 2015). QFED is based on the top-down Fire Radiative Power (FRP) approach. QFED draws on the cloud correction method used in the Global Fire Assimilation System (GFAS, Kaiser et al. 2012), but in addition employs a more sophisticated treatment of emissions from non-observed land areas (Darmenov and da Silva 2015). FRP and locations of fires are obtained from MODIS Level 2 fire and geolocation products. Level 2 fire products are gridded at $0.3125° \times 0.25°$ longitude by latitude horizontal resolution and combined to create daily mean emissions at the same resolution. QFED emissions have been independently evaluated in the WRF-Chem model in the context of the recent NASA SEAC[4]RS field campaign (Saide et al. 2015) and in the offline version of the GOCART aerosol model (Petrenko et al. 2012).

Between 1997 and 2009 we employ a scaled version of the MODIS burned-area based Global Fire Emission Dataset (GFED) version 3.1 monthly biomass burning emissions (Randerson et al. 2006; van der Werf et al. 2006). We apply biome-dependent correction factors based on the period 2003 – 2011 where both GFED v3.1 and QFEDv2.4-r6 were available. Spatially varying fractional contributions of emissions from the tropical forest, extra-tropical forest, savannah and grassland biomes were calculated using the QFED climatology. These fractional contributions were then used to stratify the monthly GFED emissions by biome. Next, for each biome, the estimated GFED emissions were used to determine a scaling factor by means of a linear regression between the globally integrated QFED and GFED emissions. The final product, the QFED-scaled GFED emission, was computed as the sum over the biomes of the scaled biome-stratified emissions.

We apply this same type of correction to the RETROv2 monthly mean emissions (Duncan et al. 2003) that cover 1980 – 1996 and are based on AVHRR, Along Track Scanning Radiometer (ATSR), and Total Ozone Mapping Spectrometer (TOMS) Aerosol Index (AI). We use the same aforementioned biome-specific scaling factors, justified by the similarities between global emissions from RETROv2 and GFEDv3.1. As can be seen in Figure 2, on monthly-mean timescales, the scaling process results in emissions that are globally consistent over time.

3. Goddard Aerosol Assimilation System

The MERRA-2 aerosol analysis uses the Goddard Aerosol Assimilation System (GAAS) that was first detailed in Buchard et al. (2015, 2016). Every 3 hours, this system assimilates quality-controlled AOD at 550 nm into the GEOS-5/GOCART modeling system. The prognostic variable in GOCART is the three-dimensional aerosol mass mixing ratio (x_i) for each species i. The forecasted column-integrated aerosol optical depth τ^f is expressed as:

$$\tau^f = \sum_{z,i} x_i \times b_{ext,i}(RH, \lambda) \times \delta z \qquad (1)$$

where δz is the atmospheric layer thickness and $b_{ext,i}(RH, \lambda)$ is the species-specific extinction coefficient at wavelength λ derived from Mie-theory for spherical particles (Wiscombe 1980) or the T-matrix approach using the updated optics for non-spherical dust as described in Meng et al. (2010). The aerosol species i include the 15 tracers previously described. The extinction coefficients ($b_{ext,i}(RH, \lambda)$) for sulfate and hydrophilic carbonaceous aerosol are a function of relative humidity (RH) following Chin et al. (2002); hygroscopic growth of sea salt aerosol follows Gerber (1985). Assumed optical properties are primarily from the Optical Properties of Aerosols and Clouds (OPAC) data set (Hess et al. 1998) with updated dust optical properties that incorporate non-sphericity (Meng et al. 2010; Colarco et al. 2014). Supplementary Table 1 provides aerosol optical properties (extinction coefficient, single scattering albedo, and asymmetry parameter) by species at 550 nm as a function of relative humidity; assumed dry sizes for each aerosol species are also given.

The AOD analysis in MERRA-2 is performed by means of the so-called *analysis splitting* method. First, a 2D analysis of AOD is performed using error covariances derived from innovation data using the maximum-likelihood method of Dee and da Silva (1999). The AOD analysis equation is:

$$\tau^a \equiv \mathrm{H}x^a = \mathrm{H}(x^f + \delta x^a)$$
$$= \tau^f + \delta\tau^a \qquad (2)$$
$$= \tau^f + \mathrm{H}\mathbf{P}^f\mathrm{H}^T(\mathrm{H}\mathbf{P}^f\mathrm{H}^T + \mathbf{R})^{-1}(\tau^o - \mathrm{H}x^f)$$

where the superscripts a, f, and o indicate the analysis, forecast (background), and observation, respectively. H is the linear observation operator that converts aerosol mass to AOD. The operators $\mathbf{P}^f$ and $\mathbf{R}$ are the background and observation error covariance matrices, respectively. The AOD analysis increments $\delta\tau^a$ are computed using a 2D version of the Physical-space Statistical Analysis System (PSAS, Cohn et al. 1998). For algorithm consistency, this analysis is performed using a natural log-transformed control variable ($\eta = \ln(AOD + \varepsilon)$ where $\varepsilon = 0.01$) as detailed in Section 3a.

Once the AOD analysis increments are obtained, the next step is to derive 3D analysis increments for the mixing ratio of each aerosol species, δx_i^a. Previous studies (e.g. Zhang et al. 2008, and references therein) have opted to simply scale the mixing ratio increments as to promote consistency with the analyzed AOD at each gridpoint, a procedure that does not make any use of error covariance information. Here we seek a relationship relating analysis increments of AOD to analysis increments of aerosol species mixing ratio that involve the corresponding error covariance operators. From the mixing ratio analysis equation implicit in Equation 2, the vertical structure of δx^a is determined by the operator $\mathbf{P}^f\mathrm{H}^T$. Therefore we seek an operator $\mathbf{Q}$ such that

$$\delta x^a = \mathbf{P}^f\mathrm{H}^T\mathbf{Q}\delta\tau^a \qquad (3)$$

Imposing the condition $\mathrm{H}\delta x^a = \delta\tau^a$ leads to

$$\mathbf{Q} = (\mathrm{H}\mathbf{P}^f\mathrm{H}^T)^{-1} \qquad (4)$$

Substituting Equation 4 into Equation 3 we arrive at:

$$\delta x^a = \mathbf{P}^f\mathrm{H}^T(\mathrm{H}\mathbf{P}^f\mathrm{H}^T)^{-1}\delta\tau^a \qquad (5)$$

Notice that the observation error covariance matrix $\mathbf{R}$ is not involved in Equation 5, and that this expression is invariant to any scaling of the background error covariance operator $\mathbf{P}^f$.

In principle, solving Equation 5 requires the 3D error covariance operator, including vertical and horizontal correlations. However, for computational reasons, we solve this equation for each vertical column separately, as the main purpose of this step is to project the horizontal AOD increments into the vertical and across species. As described in Buchard et al. (2015) and Section 3b, we employ a Local Displacement Ensemble (LDE) formulation to solve Equation 5.

a. Choice of control variable for the 2D AOD Analysis

Since AOD is not a normally distributed variable (e.g. O'Neill et al. 2000), the 2D analysis in Equation 2 is performed using the natural log-transformed AOD,

$$\eta = \ln(\tau + \varepsilon) \qquad (6)$$

as the control variable. The parameter $\varepsilon = 0.01$ is chosen as to render the distribution closest to a Gaussian using a probability plot technique (Chambers et al. 1983). Notice that for small values of τ the log-transformed variable η is linear in τ and approaches $\ln\tau$ for large AOD. This choice of control variable avoids the classical problem of log-normal distributions for small values of τ and allows for multiplicative rather than additive corrections for large τ (Henze et al. 2009; Saide et al. 2013). Notice that AOD errors τ' are related to η' errors by

$$\tau' \approx (\tau + \varepsilon)\eta' \qquad (7)$$

Therefore, even when η errors are Gaussian and flow independent, AOD errors are a sum of flow dependent ($\tau\eta'$) and flow independent ($\varepsilon\eta'$) components. To preserve linearity, analysis increments $\delta\eta^a$ are converted back to $\delta\tau^a$ before use in Equation 10.

b. Local Displacement Ensembles (LDEs)

In order to evaluate Equation 5 we employ an ensemble formulation. Let,

$$\mathbf{X} = (x_1 \quad x_2 \quad ... \quad x_E) \qquad (8)$$

where $\mathbf{X}$ is a $n_q \times n_E$ matrix (n_q is the number of aerosol concentration tracers times the number of vertical levels and n_E is the number of ensemble perturbations), for a particular column. From Equation 2 it follows that

$$\mathbf{Y} \equiv \mathrm{H}\mathbf{X} = (\mathrm{H}x_1 \quad \mathrm{H}x_2 \quad ... \quad \mathrm{H}x_E)$$
$$= (\tau_1 \quad \tau_2 \quad ... \quad \tau_E) \qquad (9)$$

Approximating the background error covariance matrix by $\mathbf{P}^f \sim \mathbf{X}\mathbf{D}\mathbf{X}^T$, where $\mathbf{D}$ is a diagonal matrix allowing for the weight of the ensemble perturbations, Equation 5 can be written as the unbiased linear regression equation:

$$\delta x^a = \mathbf{X}\mathbf{D}\mathbf{Y}^T(\mathbf{Y}\mathbf{D}\mathbf{Y}^T)^{-1}\delta\tau^a \qquad (10)$$

At this point we have made no assumption about the nature of the ensemble perturbations. MERRA-2 did not include an ensemble of aerosol forecasts, and this practical approach was developed to produce ensemble perturbations capable of producing realistic speciation and vertical structures for the mixing ratio analysis increments. The underlying assumption of our error covariance modeling exercise is that aerosol forecast errors are due primarily to

misplacements of aerosol plumes. Implicit in this assumption is that the AOD analysis (Eq. 2) removes any systematic biases. For each gridpoint, ensemble perturbations are formed by computing the difference of background aerosol mixing ratios from this central gridpoint and adjacent gridpoints within a given radius (taken as 1,000 km in MERRA-2). Ensemble perturbations are weighted according to $\exp(-4(\tau^f - \tau^a)^2)$ so that nearby gridpoints that better match the AOD analysis receive higher weights.

c. AOD Background Correction and Approximate Analysis Averaging Kernel

The MERRA-2 meteorological analysis is performed every 6 hours, while the AOD analysis occurs every 3 hours. For efficiency reasons, the overall analysis cycle in MERRA-2 is controlled by the meteorological assimilation, with two independent AOD analyses performed within each 6-hour cycle. Consider the 12Z analysis cycle. The AOD analysis is performed at 9Z and 12Z, using backgrounds that are forecasts from 6Z. The corresponding update of the GEOS-5 aerosol state occurs at 9Z and 12Z. At 12Z the proper background state should be a forecast from 9Z not from 6Z as in the (off-line) AOD analysis. Therefore a background correction is in order to account for this mismatch in background states. Denoting the previous AOD analysis and background by $\hat{\tau}^a$ and $\hat{\tau}^f$, respectively, it can be shown that

$$\tau^a = \hat{\tau}^a + (\mathbf{I} - \mathbf{A})\left(\tau^f - \hat{\tau}^f\right) \tag{11}$$

where τ^a and τ^f are the proper analysis and forecast at 12Z, and $\mathbf{A} = \mathbf{KH}$ is the analysis averaging kernel with $\mathbf{K}$ being the usual Kalman gain. For typical satellite swaths, the operator $\mathbf{A}$ evaluates to approximately zero outside the swath leading to a simple replacement of the background in those regions; elsewhere, Equation 11 provides a background correction that depends on the details of the analysis.

In practice, a diagonal approximation for the analysis averaging kernel $\mathbf{A}$ is utilized. In such an approximation, an additional AOD analysis is performed with all innovations set to 1 while preserving the actual observational coverage. This averaging kernel field is computed as the second step of each AOD analysis and is provided as an additional diagnostic for MERRA-2. Such an algorithm can be derived as a limiting case of a banded approximation for the Kalman gain $\mathbf{K}$ (derivation not shown).

d. AOD Observing System and Bias Correction

Table 2 summarizes the AOD observing system used in MERRA-2 while Figure 3 shows the total number of monthly observations by sensor for the entire reanalysis. Assimilation of satellite aerosol products requires careful data quality control and bias removal (Zhang and Reid

2006; Lary et al. 2010). MERRA-2 includes assimilation of bias-corrected AOD derived from AVHRR and MODIS radiances. Our approach involves cloud screening and homogenization of the observing system by means of a neural net retrieval (NNR) that translates cloud-cleared observed radiances into AERONET-calibrated AOD. Additional details on the NNR algorithm are given in the below, and Supplementary Table 2 provides observational error standard deviations and Kalman gains (the ratio of the background error variances to the innovation error variances) for each sensor used in the PSAS-based analysis of the log-transform AOD η. For reference, the background error decorrelation length scale is 140 km, and the background error standard deviation is 0.45. Homogeneous and isotropic covariance models are assumed, with spatially constant variances. Error covariance model parameters were estimated using maximum-likelihood method of Dee and da Silva (1999).

For the pre-EOS period and until 2002, we assimilate bias-corrected AOD derived from the 25-year record of AVHRR radiances (Heidinger et al. 2014). We discontinue AVHRR assimilation after 2002 when MODIS Aqua becomes available; both instruments have afternoon equator crossing times. Note that AVHRR only provided AOD retrievals over the oceans. After 2000, we assimilate bias-corrected AOD derived from MODIS Level 2 radiances, first from the Terra spacecraft (10:30 local solar time equator crossing) and after 2002 also from the Aqua spacecraft (13:30 local solar time equator crossing). In both cases, we use the same radiances that are provided with operational MODIS retrievals (Levy et al. 2007). Over land we use the radiances from the MODIS "Dark Target" land algorithm that are not available over bright surfaces. To include desert regions, we assimilate AOD from MISR (Kahn et al. 2005) only where the surface albedo is > 0.15. Observed Level 2 AOD from ground based AERONET stations (Holben et al. 1998) are also assimilated after 1999. Note that AERONET AOD are not reported at 550 nm; we interpolate to 550 nm using the Ångström relationship and AOD reported at adjacent channels, typically 500 and 675 nm. Because they are not currently available in Near Real Time (NRT), both MISR and AERONET observations are excluded from MERRA-2 after 2014; also, observations from these sensors are not bias corrected.

To derive 10-km resolution MODIS NNR AOD, over-ocean predictors include Level-2 multi-channel top-of-the-atmosphere (TOA) reflectances, glint, solar and sensor angles, cloud fraction (pixels are discarded when cloud fraction $> 70\%$), and albedo derived using GEOS-5 surface wind speeds. Over land, predictors are the same, except a climatological albedo is included for pixels with surface albedo < 0.15. The target of the NNR algorithm is the log-transformed AERONET AOD. For the AVHRR NNR AOD, the neural net predictors over ocean are the AVHRR Pathfinder Atmosphere-Extended (PATMOS-x)

TOA radiances at 630 and 860 nm (Heidinger et al. 2014), total precipitable water (TPW), ocean albedo (wind speed), solar and sensor angles, and the climatological GEOS-5 fractional AOD speciation. For consistency of the observing system, the target of the AVHRR NNR AOD is the MODIS NNR. As an additional quality control step, the on-line adaptive buddy check of Dee et al. (2001) is used before assimilating AOD data into GEOS-5.

To highlight the consistency imposed by the NNR algorithm on the MERRA-2 aerosol observing system, Figure 4 compares the AVHRR and MODIS Aqua NNR AOD during a month when both instruments were available. Saide et al. (2013) independently evaluated the MODIS NNR AOD against other observational datasets using the WRF-Chem data assimilation system. Post-processing techniques such as the NNR were shown to reduce biases in the assimilation relative to independent AERONET AOD observations. Additionally, compared to other bias-corrected MODIS AOD products (e.g. Zhang and Reid 2006), the NASA NNR retrieval produced a more positive impact on assimilated AOD because its less restrictive cloud fraction requirement increases data availability.

e. Control Simulation without AOD Assimilation

GEOS-5 can be run in a replay mode whereby a previous meteorological analysis, generated with an identical model version, is used to adjust the model's meteorological state (winds, temperature, specific humidity) much like a Chemical Transport Model (CTM). However, unlike a CTM, in a replay simulation the aerosol transport dynamics are entirely consistent with the model thermodynamical state at every time step between analysis updates. Using the same version of GEOS-5 as MERRA-2, we perform an EOS-period control simulation (M2REPLAY) driven by the exact MERRA-2 meteorology but without the aerosol optical depth assimilation. The aerosol fields from this control simulation are compared to those from MERRA-2 to demonstrate the impact of the AOD assimilation on aerosol fields (Section 4 and Buchard et al. (2017)).

4. Results

Figure 5a shows the timeseries of global, area-weighted monthly-mean analyzed AOD from MERRA-2. The contribution of different aerosol species to the total AOD is also shown. For comparison, we show the AOD from the previous MERRAero reanalysis (Buchard et al. 2015, 2016). Major features of the reanalysis include large increases in global AOD following major volcanic eruptions in the pre-EOS period and the seasonal cycle of AOD associated primarily with dust and biomass burning (carbonaceous) aerosols. During the EOS-period, MERRA-2 AOD is slightly higher than the AOD from MERRAero, due to higher dust and sea salt AOD that result from changes in

model physics, meteorology, and the inclusion of MISR AOD assimilation over bright desert regions in MERRA-2 (Randles et al. 2016).

a. AOD Innovation Statistics

Figure 5b shows co-located observed (O) and forecasted (F) AOD from AVHRR NNR and MODIS Aqua NNR (separately over land and ocean). Figure 5c shows the statistical relationship between the co-located assimilated observations of AOD, the forecasted AOD, and the analyzed AOD (A) for these sensors. Supplementary Figures 1-3 show similar plots for other sensors in the MERRA-2 aerosol observing system (Table 2). Additionally, Supplementary Figure 4 shows the probability distribution functions of $O - F$ and $O - A$ for each sensor in the AOD observing system. These comparisons serve as a sanity check, as it is expected that the analysis should have higher correlation and lower bias and root mean square error with respect to the assimilated observations compared to the forecast. Note that while the analysis is statistically closer to the observations than the forecast, for any given sensor, the bias is not completely eliminated due in part to the influence of other sensors on the analyzed AOD. During the EOS-period, analysis statistics are also slightly better over ocean compared to land, due to a combination of observability (fewer observations over land versus ocean) and a prevalence of aerosol source regions over land (greater aerosol type variability) (Gelaro et al. 2017; Randles et al. 2016).

Generally, forecasted AOD departs only slightly from co-located assimilated observations (Fig. 5b). The exception to this is after major volcanic eruptions (El Chichón in 1982 and Pinatubo in 1991) when the MERRA-2 forecasted AOD is significantly higher than the assimilated AVHRR observations. Buchard et al. (2017) describes the reasons for this in more detail for the Pinatubo eruption. Briefly, MERRA-2 first overestimates the volcanic plume height, injecting sulfur dioxide gas higher into the stratosphere than more recently available data suggest. This higher injection height has implications for transport and lifetime of the subsequently formed stratospheric sulfate aerosol, particularly since the assimilation does not constrain the precursor sulfur dioxide gas. Next, GOCART assumes all sulfate aerosols share the same (dry) size distribution when calculating aerosol optical properties. The sulfate aerosol size in GOCART is more appropriate for tropospheric aerosols rather than the particles that formed in the stratosphere after the Pinatubo eruption (Aquila et al. 2014). Smaller aerosol particles are more efficient at scattering light, and the underestimate in stratospheric aerosol particle size contributes to an overestimate of stratospheric and total AOD in the forecast.

b. Impact of the AOD Analysis

Figure 6 compares, for each of the satellite sensors during the EOS-period, the 3-hour forecasted and analyzed climatological AOD. We also show the climatological AOD from the replay control simulation where we have sampled M2REPLAY using the day-time Terra and Aqua orbits. In most regions, the analyzed MERRA-2 AOD resembles the forecasted AOD. Exceptions appear, for example, over China and northern Africa. Both the MERRA-2 forecasted AOD (which retains the influence of the previous assimilation step) and the analyzed AOD fields show differences relative to the control experiment. Without AOD assimilation, M2REPLAY generally has higher overall AOD, particularly over the Southern Ocean and North Atlantic, over and downwind of China, and over the Sahara.

Figure 7 shows observed, forecasted, and analyzed AOD from AVHRR for all of MERRA-2 in which we assimilate observations from this sensor and for the two years during and subsequent to the Pinatubo eruption. Here we see, as in Figure 5b and 5c, that the 3-hour forecast overestimates the observed AOD after Pinatubo. However, the analysis resembles the observations after Pinatubo. This implies a negative aerosol increment. Importantly, when a negative AOD increment is applied, it impacts all aerosol species, proportionally reducing their masses to optimally reproduce observed AOD. This can be seen in the timeseries of AOD in Figure 5a, where dust, carbonaceous, and sea salt AOD is reduced immediately after the Pinatubo eruption. This speciated AOD reduction is an artifact of applying relatively large, negative AOD increments in MERRA-2. We note that both two-dimensional total AOD increments and three-dimensional speciated aerosol mass mixing ratio increments are available as part of the MERRA-2 dataset (Global Modeling and Assimilation Office 2015g,h). Ongoing work is focused on improved forward modeling of large stratospheric volcanic eruptions to mitigate the need for large, negative AOD analysis increments (i.e. $\delta\tau_a$ in Eq. 2).

c. Independent Verification

Here we assess the performance of the analyzed AOD fields by comparing them with other, independent (non-assimilated) observations from both surface-based and aircraft sensors. In order to examine the impact of the AOD analysis, we perform these comparisons for the analyzed fields and, where available, AOD fields from the control experiment. We reiterate that it is challenging to independently validate the MERRA-2 aerosol products because most of the global, readily available spaceborne and ground-based observations are included in the assimilation. A more in-depth validation of the aerosol analysis, including other aerosol properties (vertical distribution, absorption, surface concentration) and case studies

of aerosol events is presented in Buchard et al. (2017). In this work, we focus on validation of the total column AOD only, since this is the aerosol property directly constrained in the assimilation.

1) COMPARISONS WITH AERONET

The AErosol RObotic NETwork (AERONET; http://aeronet.gsfc.nasa.gov/new_web/) is a federated global network of ground-based, automatic sunphotometers that measure direct sun and sky radiances at several wavelengths (Holben et al. 1998). AOD is obtained from direct sun measurements with an accuracy to within ±0.015. In the assimilation, we use cloud-screened Level 2.0 data (quality assured, Smirnov et al. 2000). Though AERONET is assimilated in MERRA-2 after 1999 (Table 2) and cannot be considered for independent verification, a subset of sites (∼40) have un-assimilated Level 2.0 data going back to 1993. It is instructive to examine the performance of the MERRA-2 analyzed AOD fields relative to observations at these sites. Additionally, comparing the performance of MERRA-2 AOD to M2REPLAY shows the impact of the analysis at these stations.

Table 3 compares simulated AOD to observations from long-term AERONET stations (>7 years of data), some of which also had some portion of their observations occurring prior to 1999. Statistics are computed based on hourly observations and co-located model output, and are provided both for each station's full EOS-period record (2000 – 2014), and, in parenthesis, just that part of the data record available prior to 1999. For each station, the dominant aerosol source type is indicated. For most all stations, comparing MERRA-2 to M2REPLAY reveals that the assimilation increases correlation while reducing root mean square error of the differences (RMSE) with the observed hourly AOD. Notably, even without assimilation, M2REPLAY shows considerable skill in simulating AOD at biomass burning aerosol dominated stations since emissions are observationally-based (Section 2b). For MERRA-2, comparing statistics for the period prior to 1999 and the EOS-period (within and without parenthesis, respectively), we see improved statistics once overland observations from AERONET and other EOS-sensors are incorporated into the assimilation.

Figure 8 shows the timeseries of monthly mean and standard deviation of observed AOD at three long-term AERONET sites representative of polluted continental (GSFC), biomass burning (Mongu), and dust-influenced (Capo Verde) conditions. Additional stations from Table 3 are shown in the Supplement. For each site, we also show the 30-day (centered) running mean and standard deviation of AOD from MERRA-2 and M2REPLAY. The two simulations were co-located in space and time with the corresponding hourly Level 2 AERONET observations

prior to smoothing, and statistics for all the hourly co-locations are given on each panel. MERRA-2 shows improved correlation and reduced bias relative to AERONET AOD compared to the control run. Statistics improve the least for Mongu, where M2REPLAY shows considerable skill in capturing the monthly mean AOD without aerosol assimilation (due to the use of observationally-based biomass burning emissions inventories described in Section 2b). In all cases, the variability in observed AOD is better captured by the assimilation (MERRA-2) compared to the control run (M2REPLAY).

Prior to 1999, MERRA-2 assimilates AOD from AVHRR over ocean only, and land-based observations from AERONET are assimilated starting in 1999. EOS-era satellite observations are assimilated after 2000, providing satellite observations of AOD over both land and ocean. We cannot disentangle the influence of AERONET observations alone on the AOD assimilation. However, we do see that the EOS-era observations generally better simulate AOD variability (standard deviation) compared to the impact of AVHRR ocean observations only prior to 1999. For example, prior to 1999, MERRA-2 does not generally have a standard deviation comparable to the observations in summer at GSFC (Fig. 8a), a location with AOD dominated by anthropogenic sources that have only annually varying emissions (Table 1).

The timeseries of observed and simulated AOD at Mongu (Fig. 8b) illustrates the combined impact of changing biomass burning emissions and the observing system on the analyzed AOD. Recall that prior to 1997, MERRA-2 uses emissions based on a scaled-version RETROv2 (Duncan et al. 2003) and does not include assimilated observations over land. After 1997, emissions are based on scaled-versions of GFED and QFED, both of which derive from MODIS data products. There is a clear improvement both in the monthly-mean AOD and standard deviation of AOD after 1997 when emissions change. Further improvement is seen during the EOS-period when many more observations over land become available. We note that the change from monthly-mean to daily biomass burning emissions (before and after 2010, respectively) has only a minor impact on the multi-year statistical agreement between MERRA-2 and AERONET (not shown); however, the biomass burning emissions frequency is likely more important in the context of individual events. We conclude that, importantly, timeseries analysis of MERRA-2 AOD is sensitive both to a changing observing system and changes and/or trends imposed by emission inventories.

2) COMPARISONS WITH MARITIME NETWORKS

We use historical shipborne sun photometer data to validate MERRA-2 AOD prior to the EOS-period. These data, summarized in Smirnov et al. (2002), Sakerin and Kabanov (2002) and partially used by Liu et al. (2004) for AVHRR validation, span the timeframe 1982 – 1996 and cover the cruises shown in Figure 9a. We compare the co-sampled MERRA-2 AOD at the observed wavelength closest to 550 nm (range 500 – 570 nm depending on the cruise). Observations are reported as morning or afternoon averages with an observed accuracy of ± 0.02. The model is sampled every three hours for a given observation date and then averaged 8 am – 12 pm or 12 pm – 4 pm local time for purposes of comparison. Figure 9b shows a scatter plot of the AOD comparison, with statistics reported by year. The correlation between MERRA-2 and the observations is best near dust-influenced regions (e.g. the Red Sea, Mediterranean, and near North Africa). In the remote Pacific, MERRA-2 AOD is not well-correlated with the observations, but the bias is low. Observations from a single cruise indicate that MERRA-2 AOD is biased high after the Pinatubo eruption.

During the EOS-period, the Maritime Aerosol Network (MAN; `http://aeronet.gsfc.nasa.gov/new_web/man_data.html`) employs Microtops II sun photometers aboard ships of opportunity to measure AOD. The photometers used in MAN are calibrated to have an estimated uncertainty in AOD of ± 0.02 (Smirnov et al. 2009). MAN cruises cover the period 2004 – present (Fig. 10a). These observations have not been assimilated in the MERRA-2 GAAS and therefore serve as independent validation of the assimilated AOD product. Figure 10b compares all available MAN observations with co-located MERRA-2 AOD. A high degree of correlation is found between the MERRA-2 and MAN observations, and the bias is generally low, though MERRA-2 does tend to slightly overestimate the lowest observed AOD. The impact of AOD assimilation is apparent in Figure 10c, where the M2REPLAY control simulation has lower correlation and higher bias relative to the MAN observations than MERRA-2.

3) COMPARISONS WITH AIRCRAFT OBSERVATIONS

The NASA Langley Research Center (LaRC) DIfferential Absorption Lidar (DIAL) system implements the High Spectral Resolution Lidar (HSRL) technique to retrieve aerosol extinction and AOD at 532 nm (Hair et al. 2008). The instrument also retrieves aerosol backscatter coefficients and is sensitive to polarization at three wavelengths (355, 532, and 1064 nm), measuring both above and below the aircraft (i.e. zenith and nadir). AOD are derived from nadir aerosol extinction profiles when the aircraft is above 6 km; a 1 km region below the aircraft is omitted from the AOD column as the laser and telescope are not fully aligned in this region.

During the NASA Studies of Emissions and Atmospheric Composition, Clouds, and Climate Coupling by Regional Surveys (SEAC[4]RS) campaign from August to September 2013 (`http://www.nasa.gov/mission_`

`pages/seac4rs`), the DIAL/HSRL system aboard the NASA DC8 aircraft measured AOD over a large portion of the southeastern and western continental United States, including several smoke plumes from large fires (e.g. the Rim Fire). We sample MERRA-2 along the aircraft trajectories for the entire SEAC[4]RS campaign and calculate the AOD over the same portion of the column as the observations. Figure 11a shows the performance of the MERRA-2 AOD relative to this independent observation. There is a high degree of correlation and low bias between MERRA-2 and AOD measured by the DIAL/HSRL system. As with the MAN comparison (Fig. 10), there is a high and low bias for low and high AOD, respectively. Figure 11b shows the same comparison for the control run simulation (M2REPLAY) and again demonstrates the positive impact of AOD assimilation. While we have shown the AOD comparison for the entire SEAC[4]RS campaign, Buchard et al. (2017) uses DIAL/HSRL observations to validate vertical profiles of aerosol optical properties for several aircraft campaigns and during the Yosemite Rim Fire.

The Spectrometer for Sky-Scanning, Sun-Tracking Atmospheric Research (4STAR) measured the AOD above the DC8 during the SEAC[4]RS campaign. 4STAR combines airborne sun tracking and sky scanning with spectroscopy to cover the full 350 to 1700 nm spectral range (Dunagan et al. 2013). The 4STAR AOD is computed from the observed intensity of the direct solar beam based on the Beer-Lambert Law. Shinozuka et al. (2013) give details on AOD data acquisition, screening, reduction, calibration and uncertainty analysis. During SEAC[4]RS, where ambient temperature varied more widely than in 4STAR's previous deployments, the measurements exhibited a systematic high bias in AOD. An empirical correction determined from field data and lab tests has been applied to the AOD such that, where the magnitude of the correction is comparable to the AOD itself or where the 4STAR temperature is below -15°C, AOD are flagged and not reported. This quality screening improves the accuracy of reported AOD, but also disproportionally eliminates AODs under conditions of low aerosol burden or high altitude. Comparisons with ground sun photometers before and after each flight allowed for characterization of 4STAR's optical throughput, enabling an estimate of measurement uncertainty of $\pm$ 0.02 for SEAC[4]RS.

Figure 12 compares probability distribution functions (PDFs) of co-located simulated AOD to 4STAR for both MERRA-2 and M2REPLAY for the entire SEAC[4]RS campaign. Here, we have sampled the simulations along the aircraft track, and AOD is calculated for only the portion of the atmospheric column above the aircraft. Note that differences between observed and simulated AOD at the lowest tail of the PDFs, which are representative of cleaner background air remote from major aerosol sources, are generally within the reported instrumental uncertainty. Because AOD here is not representative of the full atmospheric column, simulated AOD is sensitive to the vertical distribution of aerosol, which is not directly constrained in the assimilation. The assimilation best characterizes the mid-range of AODs ($\sim$0.13-0.21) compared to the control run. However, it underestimates the highest observed AODs (associated with biomass burning events; not shown); without assimilation (M2REPLAY), the model overestimates these high-AOD occurrences.

d. Clear-sky aerosol direct radiative effects

Atmospheric aerosols, both natural and anthropogenic, impact climate through scattering and absorption of radiation (direct radiative effect, DRE), modification of cloud microphysics (indirect effects), and thermodynamic effects (semi-direct effect of aerosol absorption). Estimating the direct radiative effect requires knowledge of the three dimensional distribution of aerosols and their optical properties. While satellites can measure AOD, a key aerosol property for determination of the DRE, cloud contamination and satellite viewing geometries can combine to produce spatial and temporal sampling biases. Furthermore, global observations of aerosol absorption optical depth (AAOD) and vertical distributions are currently even more sparsely available. Though they simulate the full aerosol life cycle without data gaps, global aerosol models are complex and produce varying estimates of DRE (Kinne et al. 2006; Schulz et al. 2006; Yu et al. 2006). In the MERRA-2 aerosol reanalysis, however, AOD is continuously available for over two decades and is optimally constrained by quality controlled satellite- and ground-based observations.

Before examining the DRE from MERRA-2, we first compare the global average AOD and AAOD to other models and reanalyses. Several forecasting centers are currently producing global aerosol reanalyses. For example, the Naval Research Laboratory (NRL) has produced an 11-year offline aerosol reanalysis at 1° resolution using the Navy Aerosol Analysis and Prediction System (NAAPS) to assimilate quality-assured and controlled MODIS Collection 5 and MISR AOD (Lynch et al. 2016). As part of the Monitoring Atmospheric Composition and Climate (MACC) project, the European Center for Medium-range Weather Forecast (ECMWF) has assimilated MODIS AOD from 2003 to the present (Morcrette et al. 2009; Benedetti et al. 2009).

Table 5 and Gelaro et al. (2017, their Fig. 14) compare the global-average AOD and AAOD from MERRA-2 to recent reanalysis estimates, including our own MERRAero offline aerosol reanalysis. Where such information is available, results are also partitioned by species and identified as either fine or coarse mode (see table footnotes for details). Also shown are the multi-model average results from Phase I of the AeroCom intercomparison project (Kinne et al. 2006), as well as both model and

observational estimates from Yu et al. (2006). Yu et al. (2006) attempt to account for satellite clear-sky biases by combining MODIS and MISR observations with the GOCART aerosol model as part of their observational estimate. Compared to MERRAero, MERRA-2 has slightly higher global average AOD due to increased contributions from dust (related to the assimilation of MISR AOD over bright surfaces) and sea salt (related to changes in model physics). MERRA-2 and NAAPS have similar global average AOD both for fine and coarse model aerosol. Global aerosol model estimates of AOD (AeroCom and Yu Model) are lower than MERRA-2, and the observational estimate (Yu Obs) is higher. The MACC reanalysis (Bellouin et al. 2013) AOD is close to the MODIS-only AOD of 0.188 (Yu et al. 2006). MACC also has more dust and sea salt AOD compared to MERRA-2, especially over the ocean (see Supplementary Table 3).

Both AOD and AAOD influence the direct impact of aerosols on the radiative energy balance of the planet. Recall that AAOD is only indirectly constrained by assimilating AOD. MERRA-2 AAOD is slightly lower than our previous aerosol reanalysis, but agrees well with the AAOD from Bellouin et al. (2013, note that the AAOD is not the MACC-native AAOD but is derived in this study). While the column-integrated AAOD from MERRA-2 generally agrees well with OMI observations in the near-UV (especially near source regions; Buchard et al. 2016), MERRA-2 overestimates absorption in regions remote from aerosol sources (e.g. the free troposphere) due to excessive amounts of black carbon aerosol in these regions (see Randles et al. 2016, for comparisons of MERRA-2 to black carbon vertical distributions from aircraft observations).

Recall that the DRE is defined as the shortwave flux difference in W m^{-2} between clear-sky (i.e. no clouds) and clear, clean-sky conditions (i.e. no aerosols or clouds). In the absence of clouds, the radiative effect of aerosols is less sensitive to the vertical distribution of aerosol absorption, although it remains sensitive to absorbing aerosols over bright surfaces (e.g. snow and deserts, Chýlek and Coakley 1974). As long-term aerosol reanalyses such as MERRA-2 continue to evolve and improve, they can potentially reduce uncertainty in the DRE, particularly once better observational constraints on aerosol absorption and vertical distribution become available and are included in the assimilation. For comparison (Fig. 13), we consider DRE estimates derived from the MACC project (Bellouin et al. 2013). DRE estimates based on observations and calculated based on the results of four global aerosol models are also considered (Yu et al. 2006).

Much of the uncertainty in the DRE reported by the Intergovernmental Panel on Climate Change (IPCC) arises from differences between estimates from global models and satellite-based estimates (Myhre 2009). Figure 13 shows the time series of clear-sky shortwave DRE from MERRA-2 and MERRAero separately over land and ocean. We also indicate DRE estimates derived from the MACC reanalysis (Bellouin et al. 2013, yellow shading) and Yu et al. (2006, both model and observational estimates, red and grey shading, respectively). Top-of-the atmosphere (TOA), surface (SFC), and atmospheric (ATM) DRE is shown, where TOA = SFC + ATM. Supplementary Table 3 presents the results of Figure 13 in tabular form and also includes a comparison of AOD, AAOD, and DRE efficiency (DRE/AOD) separately over land and ocean. Differences in the DRE efficiency highlight where model aerosol assumptions (e.g. absorption, size distribution) and environmental properties (e.g. surface albedo) contribute to discrepancies between DRE estimates from reanalyses. The DRE efficiency for MERRA-2 is within about 20% of the observational estimate (Yu et al. 2006).

As the global aerosol observing system continues to grow and provide additional information on aerosol absorption, size, type, and vertical distribution that can be assimilated, reanalysis estimates of aerosol radiative-climate effects should become more consistent with satellite-derived effects and thus reduce DRE uncertainty. For example, the GMAO is working to incorporate aerosol vertical distribution information from the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) into a future aerosol reanalysis. Unlike satellite estimates alone, reanalyses like MERRA-2 can provide detailed information on how the anthropogenic component of aerosols, and thus radiative forcing, has changed during the modern satellite era. This should lead to reduced uncertainty in assessing the human impact on climate.

5. Summary and Conclusions

This paper describes the NASA GMAO's MERRA-2 aerosol reanalysis, the first satellite era (1980 – onward) reanalysis in which both meteorological and aerosol observations are jointly assimilated. We use the GOCART aerosol module coupled to the GEOS-5 Data Assimilation System (DAS) to include radiatively-active prognostic aerosol mass tracers and the Goddard Aerosol Assimilation System (GAAS) to assimilate bias-corrected AOD from AVHRR and MODIS (Terra and Aqua), MISR AOD over desert regions, and ground-based AERONET AOD. Publicly available gridded output is available at hourly, three-hourly, and monthly timescales for both column-integrated and three-dimensionally resolved aerosol mass, optical properties, and other aerosol diagnostics. These data are available through the GES DISC at `http://disc.sci.gsfc.nasa.gov/mdisc/`.

In this paper we have provided information about the MERRA-2 Data Assimilation System relevant for the aerosol assimilation, including a description of the aerosol module and emissions inventories. We describe both the process of AOD assimilation and the necessary prior data

quality control. In MERRA-2 the only property directly constrained by the assimilation is the AOD. Other aerosol properties such as aerosol speciation and vertical distribution, all of which are available, are indirectly constrained by the assimilation. Here, we focus our evaluation and validation efforts on the AOD only. Our companion study (Buchard et al. 2017) presents an evaluation and validation of aerosol properties that are indirectly impacted by the MERRA-2 AOD assimilation. In both studies, to examine the overall impact of the AOD assimilation, we perform an EOS-period control simulation in replay mode where the same version of the model is driven by MERRA-2 analyzed meteorology, but without AOD assimilation.

We begin by assessing innovation statistics as a sanity check, as we expect analyzed AOD fields to better match assimilated AOD than the forecast AOD. As shown by Lynch et al. (2016), a well-performing forward model is as equally important as the AOD assimilation process itself. During the EOS-period, a period for which the GEOS-5/GOCART system has been extensively evaluated and refined, we show that in most regions, there is little difference between the 3-hour forecast and analyzed AOD fields. Thus, on average, the assimilation process needs to apply relatively small AOD analysis increments ($\delta\tau^a$) to agree with the assimilated observations. This is an indication that the model is relatively unbiased with respect to the AOD measurements it assimilates.

The forward modeling system in MERRA-2, however, has not been optimized to deal with the larger, stratospheric aerosol produced from SO_2 oxidation after major volcanic eruptions. We are currently working on improving the representation of stratospheric SO_2 injection by past volcanoes, in terms of plume height and injection magnitude by incorporating data from Carn et al. (2016). This dataset combines UV measurements from TOMS, OMI, and OMPS with infrared data from TOVS, AIRS and IASI to produce a database of volcanic SO_2 injection amount and plume top height from 1979 to present. Future inclusion of a separate volcanic aerosol tracer or aerosol microphysics will improve the representation of stratospheric aerosol size. Improvements in the forecasted AOD after major eruptions will mitigate the need for large, negative AOD increments, and we expect less of an artificial impact on the timeseries of AOD (and DRE) for individual aerosol species.

Next we focus on evaluating and validating the MERRA-2 analyzed AOD fields. First, we compare hourly AOD to AERONET observations, which were not assimilated until 1999. We find that at all stations, the assimilated AOD better matches the observations than the control simulation, not only in terms of correlation but also in terms of more realistic variability and reduced root mean squared error. The assimilation produces the least impact for biomass burning regions, where emissions during the EOS-period have been well-tuned to MODIS data

products. Comparisons to AERONET also reveal that both changes in the aerosol observing system between the pre-EOS and EOS periods and changes in prescribed aerosol emissions inventories (especially for biomass burning) impact MERRA-2 AOD.

Shipborne sunphotometer data is used for independent validation of MERRA-2 AOD both for the pre-EOS and EOS periods. Compared to historic AOD cruise measurements, we find that MERRA-2 correlates reasonably well with the observations reported as morning or afternoon averages. The bias between analyzed and observed AOD is generally within the instrumental uncertainty. In part because they are recorded with time precision that allows for unambiguous sampling of the model, present-day Maritime Aerosol Network observations show even greater agreement with analyzed AOD fields. Furthermore, there is a clear, positive impact of the assimilation on the AOD as compared to the control simulation.

We conclude our model AOD validation using aircraft observations taken during the recent NASA SEAC[4]RS campaign over the southeastern and western United States in August-September, 2013. First, we compare to AOD retrievals from the DIAL/HSRL instrument. Considering that the model is spatially coarse compared to the high-resolution observations, both the MERRA-2 AOD and AOD from the control simulation well-simulate observed AOD, with improvements seen for the analyzed AOD fields. Similarly, comparisons between these two simulations and the 4STAR instrument, which measured AOD above the aircraft during SEAC[4]RS, reveal that the AOD assimilation tends to increase the lowest-simulated and mid-range AOD while decreasing the highest-simulated AOD during the campaign. Overall, the resulting campaign-wide AOD distribution from MERRA-2 better matches the observed 4STAR distribution. Notably, this improved agreement occurs despite the fact that the above-aircraft AOD is a function of the aerosol vertical distribution, an aerosol property only indirectly impacted by the assimilation (Buchard et al. 2017).

Finally, we compare MERRA-2 AOD, AAOD, and the clear-sky shortwave Direct Radiative Effect (DRE) amongst recent aerosol reanalyses, models, and observations. MERRA-2 global average AOD is higher than the model-simulated AOD, but AAOD from MERRA-2 is lower than observation studies suggest. Estimates of clear-sky DRE from MERRA-2 are closer to observationally-based estimates than models, but examination of the DRE efficiency reveals a potential need to revise aerosol optical property assumptions (e.g. absorption, size distribution). We caution that in all-sky (cloudy) conditions, the DRE of aerosols from MERRA-2 will be highly sensitive to the vertical distribution of aerosol absorption (Chýlek and Coakley 1974). While the overall MERRA-2 vertical profile of aerosol extinction is reasonable compared to observations (Buchard et al. 2016), black carbon

aerosol and its associated absorption are overestimated at higher altitudes, particularly away from source regions (e.g. compared to HIPPO observations from Schwarz et al. (2013) as shown in Randles et al. 2016). Furthermore, anthropogenic aerosol pre-industrial to present-day radiative forcing (RF), which we do not consider in this study, depends both on aerosol distributions and aerosol speciation, the latter of which is strongly controlled by lower boundary conditions (i.e. emissions inventories) rather than the assimilation of AOD. Both the AOD and DRE are impacted by a changing observing system and, in some cases, changing emissions inventories. Nevertheless, long-term aerosol reanalyses like MERRA-2 have the potential to reduce uncertainty in estimates of aerosol-radiation effects, particularly as they are further constrained by additional aerosol observations (e.g. vertical distribution, absorption, multi-wavelength information to distinguish aerosol types).

As demonstrated in here, in Randles et al. (2016), and in Buchard et al. (2017), in many cases MERRA-2 shows considerable skill in simulating numerous observable aerosol properties. However, we emphasize that only the 550 nm AOD is constrained in the reanalysis, and it is only constrained *when* and *where* data are available. Prior to the EOS-period, observations were primarily from AVHRR over ocean only. While data volume, especially over land, increased markedly after 2000, AOD observations are only available for the sunlit portion of the globe, depend strongly on satellite viewing geometry, and are subject to meteorological conditions (e.g. cloudiness). Without available data to assimilate, the assimilation naturally draws toward the forecast from the GEOS-5/GOCART model. Care must always be taken when considering aerosol products indirectly constrained by the assimilation, and trends in reanalysis aerosol properties including AOD must be considered within the context of a changing observing system and a forecast influenced by prescribed aerosol emissions inventories.

The MERRA-2 aerosol reanalysis is a major step towards an Integrated Earth Systems Analysis that will one day incorporate atmospheric, constituent, oceanic, and land observations to provide a scientific, internally-consistent gridded description of the state of the Earth system and how it is evolving over the satellite era. Future work on the aerosol analysis will focus on improving the aerosol forecast in order to minimize the needed analysis increments. We are also actively working on incorporating observations from additional space- and ground-based sensors from an ever-expanding aerosol observation network, particularly observations that provide additional infomation content (e.g. multi-spectral measurements and vertical structure from space-based lidar).

References

Aquila, V., C. I. Garfinkel, P. Newman, L. Oman, and D. Waugh, 2014: Modifications of the Quasi-Biennial Oscillation by a geoengineering perturbation of the stratospheric aerosol layer. *Geophysical Research Letters*, **41 (5)**, 1738–1744, doi:10.1002/2013GL058818.

Bellouin, N., J. Quaas, J.-J. Morcrette, and O. Boucher, 2013: Estimates of aerosol radiative forcing from the MACC reanalysis. *Atmospheric Chemistry and Physics*, **13 (4)**, 2045–2062, doi:10.5194/acp-13-2045-2013.

Benedetti, A., and Coauthors, 2009: Aerosol analysis and forecast in the European Centre for Medium-range Weather Forecasts Integrated Forecast System: 2. Data assimilation. *Journal of Geophysical Research: Atmospheres*, **114 (D13)**, doi:10.1029/2008JD011115.

Bian, H., and Coauthors, 2013: Source attributions of pollution to the eWestern Arctic during the NASA ARCTAS field campaign. *Atmospheric Chemistry and Physics*, **13 (9)**, 4707–4721, doi:10.5194/acp-13-4707-2013.

Bloom, S., L. Takacs, A. DaSilva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev*, **124**, 1256–1271.

Bocquet, M., and Coauthors, 2015: Data assimilation in atmospheric chemistry models: Current status and future prospects for coupled chemistry meteorology models. *Atmospheric Chemistry and Physics*, **15 (10)**, 5325–5358, doi:10.5194/acp-15-5325-2015.

Bosilovich, M. G., and Coauthors, 2016: MERRA-2: Initial evaluation of the climate. Technical Report Series on Global Modeling and Data Assimilation NASA/TM-2015-104606 43, NASA Global Modeling and Assimilation Office. URL https://gmao.gsfc.nasa.gov/pubs/tm/docs/Bosilovich803.pdf.

Buchard, V., and Coauthors, 2014: Evaluation of GEOS-5 sulfur dioxide simulations during the Frostburg, MD 2010 field campaign. *Atmospheric Chemistry and Physics*, **14 (4)**, 1929–1941, doi:10.5194/acp-14-1929-2014.

Buchard, V., and Coauthors, 2015: Using the OMI aerosol index and absorption aerosol optical depth to evaluate the NASA MERRA Aerosol Reanalysis. *Atmospheric Chemistry and Physics*, **15 (10)**, 5743–5760, doi:10.5194/acp-15-5743-2015.

Buchard, V., and Coauthors, 2016: Evaluation of the surface $PM_{2.5}$ in Version 1 of the NASA MERRA Aerosol Reanalysis over the

United States. *Atmospheric Environment*, **125**, 100–111, doi:10. 1016/j.atmosenv.2015.11.004.

Buchard, V., and Coauthors, 2017: The MERRA-2 aerosol reanalysis, 1980 - onward, Part 2: Evaluation and case studies. *Journal of Climate*, **accepted**.

Carn, S. A., L. Clarisse, and A. J. Prata, 2016: Multi-decadal satellite measurements of global volcanic degassing. *Journal of Volcanology and Geothermal Research*, **31**, 99–134, doi:10.1016/j.volgeores. 2016.01.002.

Chambers, J., W. Cleveland, B. Kleiner, and P. Tukey, 1983: *Graphical Methods for Data Analysis*. Wadsworth.

Chin, M., and Coauthors, 2002: Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and sun photometer measurements. *Journal of the Atmospheric Sciences*, **59**, 461–483, doi:10.1175/1520-0469(2002)059⟨0461:TAOTFT⟩2.0. CO;2.

Chýlek, P., and J. A. Coakley, 1974: Aerosol and climate. *Science*, **183**, 75–77.

Cohn, S., A. da Silva, and J. Guo, 1998: Assessing the effects of data selection with the dao physical-space statistical analysis system. *Monthly Weather Review*, **126**, 2913–2926.

Colarco, P., A. da Silva, M. Chin, and T. Diehl, 2010: Online simulations of global aerosol distributions in the NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth. *Journal of Geophysical Research*, **115 (D14207)**, doi: 10.1029/2009JD012820.

Colarco, P. R., E. P. Nowottnick, C. A. Randles, B. Yi, P. Yang, K.-M. Kim, J. A. Smith, and C. G. Bardeen, 2014: Impact of radiatively interactive dust aerosols in the NASA GEOS-5 climate model: Sensitivity to dust particle shape and refractive index. *Journal of Geophysical Research: Atmospheres*, **119 (2)**, 753–786, doi: 10.1002/2013JD020046.

Darmenov, A. S., and A. da Silva, 2015: The Quick Fire Emissions Dataset (QFED) - Documentation of versions 2.1, 2.2 and 2.4. Technical Report Series on Global Modeling and Data Assimilation, NASA//TM-2015-104606, NASA Global Modeling and Assimilation Office. URL https://gmao.gsfc.nasa.gov/pubs/ docs/Darmenov796.pdf.

Dee, D. P., and A. M. da Silva, 1999: Maximum-likelihood estimation of forecast and observation error covariance parameters. Part I: Methodology. *Monthly Weather Review*, **127**, 1811–1834.

Dee, D. P., L. Rukhovets, R. Todling, A. M. da Silva, and L. J. W., 2001: An adaptive buddy check for observational quality control. *Quarterly Journal of the Royal Meteorological Society*, **127**, 2451–2471, doi:10.1002/qj.49712757714.

Diehl, T., A. Heil, M. Chin, X. Pan, D. Streets, M. Schultz, and S. Kinne, 2012: Anthropogenic, biomass burning, and volcanic emissions of black carbon, organic carbon, and SO_2 from 1980 to 2010 for hindcast model experiments. *Atmospheric Chemistry and Physics Discussions*, **12 (9)**, 24 895–24 954, doi:10.5194/ acpd-12-24895-2012.

Dunagan, S. E., and Coauthors, 2013: Spectrometer for Sky-Scanning Sun-Tracking Atmospheric Research (4STAR): Instrument technology. *Remote Sensing*, **5 (8)**, 3872, doi:10.3390/rs5083872.

Duncan, B. N., R. V. Martin, A. C. Staudt, R. Yevich, and J. A. Logan, 2003: Interannual and seasonal variability of biomass burning emissions constrained by satellite observations. *Journal of Geophysical Research: Atmospheres*, **108 (D2)**, doi:10.1029/2002JD002378.

European Comission, 2011: European Commission/Joint Research Centre (JRC)/Netherlands Environmental Assessment agency (PBL): Emission Database for Global Atmospheric Research (EDGAR), release version 4.2. URL http://edgar.jrc.ec.europa.eu.

European Commission, 2010: European Commission/Joint Research Centre (JRC)/Netherlands Environmental Assessment agency (PBL): Emission Database for Global Atmospheric Research (EDGAR), release version 4.1. URL http://edgar.jrc.ec.europa.eu.

Eyring, V., H. W. Köhler, J. van Aardenne, and A. Lauer, 2005: Emissions from international shipping: 1. The last 50 years. *Journal of Geophysical Research: Atmospheres*, **110 (D17)**, doi:10.1029/ 2004JD005619.

Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, **accepted**.

Gerber, H. E., 1985: Relative-humidity parameterization of the Navy Aerosol Model (NAM). Tech. Rep. NTIS ADA1632090, Naval Research Laboratory, Washington, DC.

Ginoux, P., M. Chin, I. Tegen, J. M. Prospero, B. Holben, O. Dubovik, and S.-J. Lin, 2001: Sources and distributions of dust aerosols simulated with the GOCART model. *Journal of Geophysical Research: Atmospheres*, **106 (D17)**, 20 255–20 273, doi:10.1029/ 2000JD000053.

Giordano, L., and Coauthors, 2015: Assessment of the MACC reanalysis and its influence as chemical boundary conditions for regional air quality modeling in AQMEII-2. *Atmospheric Environment*, **115**, 371 − 388, doi:10.1016/j.atmosenv.2015.02.034.

Global Modeling and Assimilation Office, 2015a: `inst3_3d_aer_Nv`: MERRA-2 3D Aerosol mass mixing ratios. Instantaneous, 3-hourly (model-level, 72 eta levels), version 5.12.4, doi:10.5067/ LTVB4GPCOTK2.

Global Modeling and Assimilation Office, 2015b: `tavg1_2d_aer_Nx`: MERRA-2 2D Aerosol diagnostics. Time-averaged, hourly, version 5.12.4, doi:10.5067/KLICLTZ8EM9D.

Global Modeling and Assimilation Office, 2015c: `tavg1_2d_adgr_Nx`: MERRA-2 2D Extended aerosol diagnostics. Time-averaged, 3-hourly (column-integrated or surface), version 5.12.4, doi:10.5067/Q9QMY5PBNV1T.

Global Modeling and Assimilation Office, 2015d: `inst3_3d_asm_Nv`: MERRA-2 3D Assimilated meteorological fields. Instantaneous, 3-hourly (model-level, 72 eta levels), version 5.12.4, doi:10.5067/ WWQSXQ8IVFW8.

Global Modeling and Assimilation Office, 2015e: `tavg3_3d_asm_Nv`: MERRA-2 3D Assimilated meteorological fields. Time-averaged, 3-hourly (model-level, 72-eta levels), version 5.12.4, doi:10.5067/ SUOQESM06LPK.

Global Modeling and Assimilation Office, 2015f: `tavg1_2d_rad_Nx`: MERRA-2 Radiation diagnostics. Time-averaged, 3-hourly (single-level), version 5.12.4, doi:10.5067/HNGA0EWW0R09.

Global Modeling and Assimilation Office, 2015g: `inst3_3d_gas_Nv`: MERRA-2 3D Aerosol mass mixing ratio analysis increments. Instantaneous, 3-hourly (model-level, 72 eta levels), version 5.12.4, doi:10.5067/96BUID8HGGX5.

Global Modeling and Assimilation Office, 2015h: `inst3_2d_gas_Nx`: MERRA-2 2D Aerosol optical depth analysis increments. Instantaneous 3-hourly (column-integrated), version 5.12.4, doi:10.5067/HM00OHQBHKTP.

Gong, S. L., 2003: A parameterization of sea-salt aerosol source function for sub- and super-micron particles. *Global Biogeochemical Cycles*, **17 (4)**, doi:10.1029/2003GB002079.

Guenther, A., and Coauthors, 1995: A global model of natural volatile organic compound emissions. *Journal of Geophysical Research: Atmospheres*, **100 (D5)**, 8873–8892, doi:10.1029/94JD02950.

Hair, J. W., and Coauthors, 2008: Airborne High Spectral Resolution Lidar for profiling aerosol optical properties. *Applied Optics*, **47 (36)**, 6734–6752, doi:10.1364/AO.47.006734.

Heidinger, A. K., M. J. Foster, A. Walther, and X. Zhao, 2014: The Pathfinder Atmospheres – Extended AVHRR climate dataset. *Bulletin of the American Meteorological Society*, 909–922, doi:10.1175/BAMS-D-12-00246.1.

Henze, D. K., J. H. Seinfeld, and D. T. Shindell, 2009: Inverse modeling and mapping US air quality influences of inorganic $PM_{2.5}$ precursor emissions using the adjoint of GEOS-Chem. *Atmospheric Chemistry and Physics*, **9 (16)**, 5877–5903, doi:10.5194/acp-9-5877-2009.

Hess, M., P. Koepke, and I. Schult, 1998: Optical Properties of Aerosols and Clouds: The software package OPAC. *Bulletin of the American Meteorological Society*, **79**, 831–844, doi:10.1175/1520-0477(1998)079⟨0831:OPOAAC⟩2.0.CO;2.

Holben, B., and Coauthors, 1998: AERONET – A federated instrument network and data archive for aerosol characterization. *Remote Sensing of Environment*, **66 (1)**, 1–16, doi:10.1016/S0034-4257(98)00031-5.

Inness, A., and Coauthors, 2013: The MACC reanalysis: An 8 yr data set of atmospheric composition. *Atmospheric Chemistry and Physics*, **13 (8)**, 4073–4109, doi:10.5194/acp-13-4073-2013.

Jaeglé, L., P. K. Quinn, T. S. Bates, B. Alexander, and J.-T. Lin, 2011: Global distribution of sea salt aerosols: New constraints from in situ and remote sensing observations. *Atmospheric Chemistry and Physics*, **11 (7)**, 3137–3157, doi:10.5194/acp-11-3137-2011.

Kahn, R. A., B. J. Gaitley, J. V. Martonchik, D. J. Diner, K. A. Crean, and B. Holben, 2005: Multiangle Imaging Spectroradiometer (MISR) global aerosol optical depth validation based on 2 years of coincident Aerosol Robotic Network (AERONET) observations. *Journal of Geophysical Research: Atmospheres*, **110 (D10)**, doi:10.1029/2004JD004706.

Kaiser, J. W., and Coauthors, 2012: Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power. *Biogeosciences*, **9 (1)**, 527–554, doi:10.5194/bg-9-527-2012.

Kessner, A. L., J. Want, R. C. Levy, and P. R. Colarco, 2013: Remote sensing of surface visibility from space: A look at the United States East Coast. *Atmospheric Environment*, **81 (C)**, 136–147, doi:10.1016/j.atmosenv.2013.08.050.

Kinne, S., and Coauthors, 2006: An AeroCom initial assessment – Optical properties in aerosol component modules of global models. *Atmospheric Chemistry and Physics*, **6 (7)**, 1815–1834, doi:10.5194/acp-6-1815-2006.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, , and S. Lord, 2009: Introduction of the GSI into the NCEP Global Data Assimilation System. *Weather Forecasting*, **24**, 1691–1705.

Lana, A., and Coauthors, 2011: An updated climatology of surface dimethlysulfide concentrations and emission fluxes in the global ocean. *Global Biogeochemical Cycles*, **25 (1)**, doi:10.1029/2010GB003850.

Lary, D., L. A. Remer, D. MacNeil, B. Roscoe, and S. Paradise, 2010: Machine learning and bias correction of MODIS aerosol optical depth. *IEEE Geosci. Remote Sens. Lett.*, **6 (694)**, doi:10.1109/LGRS.2009.2023605.

Levy, R. C., L. A. Remer, S. Mattoo, E. F. Vermote, and Y. J. Kaufman, 2007: Second-generation operational algorithm: Retrieval of aerosol properties over land from inversion of Moderate Resolution Imaging Spectroradiometer spectral reflectance. *Journal of Geophysical Research: Atmospheres*, **112 (D13)**, doi:10.1029/2006JD007811.

Lin, S. J., 2004: A vertically Lagrangian finite-volume dynamical core for global models. *Mon. Wea. Rev.*, **132**, 2293–2307.

Liu, L., M. I. Mishchenko, I. Geogdzhayev, A. Smirnov, S. M. Sakerin, D. M. Kabanov, and O. A. Ershov, 2004: Global validation of two-channel AVHRR aerosol optical thickness retrievals over the oceans. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **88 (1-3)**, 97–109, doi:10.1016/j.jqsrt.2004.03.031.

Lynch, P., and Coauthors, 2016: An 11-year global gridded aerosol optical thickness reanalysis (v1.0) for atmospheric and climate sciences. *Geoscientific Model Development*, **9**, 1489–1522, doi:10.5194/gmd-9-1489-2016.

Marticorena, B., and G. Bergametti, 1995: Modeling the atmospheric dust cycle: 1. Design of a soil-derived dust emission scheme. *Journal of Geophysical Research: Atmospheres*, **100 (D8)**, 16 415–16 430, doi:10.1029/95JD00690.

McCarty, W., L. Coy, R. Gelaro, A. Huang, D. Merkova, E. B. Smith, M. Sienkiewicz, and K. Wargan, 2016: MERRA-2 input observations: Summary and Assessment. Technical Report Series on Global Modeling and Data Assimilation, NASA TM-2016-104606 46, NASA Global Modeling and Assimilation Office. URL https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/docs/.

Meng, Z., P. Yang, G. W. Kattawar, L. Bi, K. N. Liou, and I. Laszlo, 2010: Single-scattering properties of tri-axial ellipsoidal mineral dust aerosols: A database for application to radiative transfer calculations. *J Aerosol Sci*, **41 (5)**, 501–512, doi:10.1016/j.jaerosci.2010.02.00.

Molod, A., L. Takacs, M. S. andJ. Bacmeister, I.-S. Song, and A. Eichmann, 2012: The GEOS-5 atmospheric general circulation model: Mean climate and development from MERRA to Fortuna. Technical Report Series on Global Modeling and Data Assimilation NASA/TM-2012-104606, NASA Global Modeling and Assimilation Office, 117 pp. URL https://gmao.gsfc.nasa.gov/pubs/docs/Molod484.pdf.

Molod, A. M., L. L. Takas, M. Suarez, and J. Bacmeister, 2015: Development of the GEOS-5 atmospheric general circulation model: Evolution from MERRA to MERRA-2. *Geoscientific Model Development*, **8**, 1339–1356, doi:10.5194/gmd-8-1339-2015.

Morcrette, J.-J., and Coauthors, 2009: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: Forward modeling. *Journal of Geophysical Research*, **114 (D06206)**, doi:10.1029/2008JD011235.

Myhre, G., 2009: Consistency between satellite-derived and modeled estimates of the Direct Aerosol Effect. *Science*, **325**, 187–190, doi:10.1126/science.117446.

Nowottnick, E., P. Colarco, A. da Silva, D. Hlavka, and M. McGill, 2011: The fate of the Saharan dust across the Atlantic and implications for a Central Amercian dust barrier. *Atmospheric Chemistry and Physics*, **11 (16)**, 8415–8431, doi:10.5194/acp-11-8415-2011.

Nowottnick, E., P. Colarco, R. Ferrare, G. Chen, S. Ismail, B. Anderson, and E. Browell, 2010: Online simulations of mineral dust aerosol distributions: Comparisons to NAMMA observations and sensitivity to dust emission parameterization. *Journal of Geophysical Research: Atmospheres*, **115 (D3)**, doi:10.1029/2009JD012692.

O'Neill, N., A. Ignatov, B. N. Holben, and T. F. Eck, 2000: The lognormal distribution as a reference for reporting aerosol optical depth statistics: Empirical tests using multi-year, multi-site AERONET Sunphotometer data. *Geophysical Research Letters*, **27 (20)**, 3333–3336.

Petrenko, M., R. Kahn, M. Chin, A. Soja, T. Kucsera, and Harshvardhan, 2012: The use of satellite-measured aerosol optical depth to constrain biomass burning emissions source strength in the global model GOCART. *Journal of Geophysical Research: Atmospheres*, **117 (D18)**, doi:10.1029/2012JD017870, d18212.

Putman, W., and S.-J. Lin, 2007: Finite volume transport on various cubed sphere grids. *J. Comput. Phys.*, **227**, 5578, doi:10.1016/j.jcp.2007.07.022.

Randerson, J. T., and Coauthors, 2006: The impact of boreal forest fire on climate warming. *Science*, **314 (5802)**, 1130–1132, doi:10.1126/science.1132075.

Randles, C. A., and Coauthors, 2016: The MERRA-2 Aerosol Assimilation. Technical Report Series on Global Modeling and Data Assimilation, NASA TM-2016-104606 45, NASA Global Modeling and Assimilation Office. URL https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/docs/.

Reale, O., K. M. Lau, A. da Silva, and T. Matsui, 2014: Impact of assimilated and interactive aerosol on tropical cyclogenesis. *Geophysical Research Letters*, **41 (9)**, 3282–3288, doi:10.1002/2014GL059918.

Reichle, R. H., and Q. Liu, 2014: Observation-corrected precipitation estimates in GEOS-5. Tech. Rep. 35, NASA Global Modeling and Assimilation Office. URL https://gmao.gsfc.nasa.gov/pubs/docs/Reichle734.pdf.

Reichle, R. H., Q. Liu, R. D. Koster, C. S. Draper, S. P. P. Mahanama, and G. S. Partyka, 2017: Land surface precipitation in MERRA-2. *Journal of Climate*, **accepted**.

Rienecker, M. M., and Coauthors, 2008: The GEOS-5 Data Assimilation System – Documentation of versions 5.0.1, 5.1.0, and 5.2.0. Tech. Rep. 27, NASA Global Modeling and Assimilation Office. URL https://gmao.gsfc.nasa.gov/pubs/docs/Rienecker369.pdf.

Rienecker, M. M., and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective analysis for Research and Applications. *Journal of Climate*, **24**, 3624–3648, doi:10.1175/JCLI-D-11-00015.1.

Saide, P. E., G. R. Carmichael, Z. Liu, C. S. Schwartz, H. C. Lin, A. M. da Silva, and E. Hyer, 2013: Aerosol optical depth assimilation for a size-resolved sectional model: Impacts of observationally constrained, multi-wavelength and fine mode retrievals on regional scale analyses and forecasts. *Atmospheric Chemistry and Physics*, **13**, 10 425–10 444, doi:doi:10.5194/acp-13-10425-2013.

Saide, P. E., and Coauthors, 2015: Revealing important nocturnal and day-to-day variations in fire smoke emissions through a multiplatform inversion. *Geophysical Research Letters*, **42 (9)**, 3609–3618, doi:10.1002/2015GL063737.

Sakerin, S. M., and D. M. Kabanov, 2002: Spatial inhomogeneities and the spectral behavior of atmospheric aerosol optical depth over the Atlantic Ocean. *Journal of the Atmospheric Sciences*, **59**, 484–500, doi:10.1175/1520-0469(2002)059⟨0484:SIATSB⟩2.0.CO;2.

Schulz, M., and Coauthors, 2006: Radiative forcing by aerosols as derived from the AeroCom present-day and pre-industrial simulations. *Atmospheric Chemistry and Physics*, **6**, 5225–5246, doi:10.5194/acp-6-5225-2006.

Schutgens, N. A. J., T. Miyoshi, T. Takemura, and T. Nakajima, 2010: Applying an ensemble Kalman filter to the assimilation of AERONET observations in a global aerosol transport model. *Atmospheric Chemistry and Physics*, **10 (5)**, 2561–2576, doi:10.5194/acp-10-2561-2010.

Schwarz, J. P., and Coauthors, 2013: Global-scale seasonally resolved black carbon vertical profiles over the Pacific. *Geophysical Research Letters*, **40 (20)**, 5542–5547, doi:10.1002/2013GL057775.

Sekiyama, T. T., T. Y. Tanaka, A. Shimizu, and T. Miyoshi, 2010: Data assimilation of CALIPSO aerosol observations. *Atmospheric Chemistry and Physics*, **10 (1)**, 39–49, doi:10.5194/acp-10-39-2010.

Shi, Y., J. Zhang, J. S. Reid, B. Holben, E. J. Hyer, and C. Curtis, 2011: An analysis of the collection 5 MODIS over-ocean aerosol optical depth product for its implication in aerosol assimilation. *Atmospheric Chemistry and Physics*, **11 (2)**, 557–565, doi:10.5194/acp-11-557-2011.

Shinozuka, Y., and Coauthors, 2013: Hyperspectral aerosol optical depths from TCAP flights. *J. Geophys. Res.*, **118**, 1–15, doi:10.1002/2013jd020596.

Smirnov, A., B. Holben, T. Eck, and I. Slutsker, 2000: Cloud-screening and quality control algorithms for the AERONET database. *Remote Sensing of Environment*, **73 (3)**, 337 – 349, doi:10.1016/S0034-4257(00)00109-7.

Smirnov, A., B. N. Holben, Y. J. Kaufman, O. Dubovik, T. F. Eck, I. Slutsker, C. Pietras, and R. N. Halthore, 2002: Optical properties of atmospheric aerosol in maritime environments. *Journal of the Atmospheric Sciences*, **59**, 501–523, doi:10.1175/1520-0469(2002)059⟨0501:OPOAAI⟩2.0.CO;2.

Smirnov, A., and Coauthors, 2009: Maritime Aerosol Network as a component of Aerosol Robotic Network. *Journal of Geophysical Research: Atmospheres*, **114 (D06204)**, doi:10.1029/2008JD011257.

Takacs, L. L., M. J. Suarez, and R. Todling, 2015: Maintaining atmospheric mass and water balance within reanalysis. Technical Report Series on Global Modeling and Data Assimilation, NASA/TM-2014-104606 37, NASA Global Modeling and Assimilation Office. URL http://gmao.gsfc.nasa.gov/pubs/docs/Takacs737.pdf.

Textor, C., and Coauthors, 2006: Analysis and quantification of the diversities of aerosol life cycles within AeroCom. *Atmospheric Chemistry and Physics*, **6 (7)**, 1777–1813, doi:10.5194/acp-6-1777-2006.

van der Werf, G. R., J. T. Randerson, L. Giglio, G. J. Collatz, P. S. Kasibhatla, and A. F. Arellano Jr., 2006: Interannual variability in global biomass burning emissions from 1997 to 2004. *Atmospheric Chemistry and Physics*, **6 (11)**, 3423–3441, doi:10.5194/acp-6-3423-2006.

Wiscombe, W. J., 1980: Improved Mie scattering algorithms. *Appl. Opt.*, **19 (9)**, 1505–1509, doi:10.1364/AO.19.001505.

Wu, W.-S., R. Purser, and D. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905–2916.

Yu, H., and Coauthors, 2006: A review of measurement-based assessments of the aerosol direct radiative effect and forcing. *Atmospheric Chemistry and Physics*, **6 (3)**, 613–666, doi:10.5194/acp-6-613-2006.

Zhang, J., and J. S. Reid, 2006: MODIS aerosol product analysis for data assimilation: Assessment of over-ocean Level 2 aerosol optical thickness retrievals. *Journal of Geophysical Research: Atmospheres*, **111 (D22)**, doi:10.1029/2005JD006898.

Zhang, J., and J. S. Reid, 2009: An analysis of clear sky and contextual biases using an operational over ocean MODIS aerosol product. *Geophysical Research Letters*, **36 (15)**, doi:10.1029/2009GL038723.

Zhang, J., J. S. Reid, D. L. Westphal, N. L. Baker, and E. J. Hyer, 2008: A system for operational aerosol optical depth data assimilation over global oceans. *Journal of Geophysical Research: Atmospheres*, **113 (D10)**, doi:10.1029/2007JD009065, d10208.

Zhang, Y., M. Bocquet, V. Mallet, C. Seigneur, and A. Baklanov, 2012: Real-time air quality forecasting, Part I: History, techniques, and current status. *Atmospheric Environment*, **60**, 632 – 655, doi:10.1016/j.atmosenv.2012.06.031.

TABLE 1. Aerosol and Precursor Emissions in MERRA-2[a]

Aerosol Type	Source	Description[b,c]
Dust	Wind-driven	Static topographic depression source map ($0.3125° \times 0.25°$, Ginoux et al. 2001)
Sea Salt	Wind-driven	see Section 2b
Volcanic SO_2	AeroCom Phase II (HCA0 v2; Diehl et al. 2012)	Daily degassing and eruptive volcanos (1980 – 2010)
Biogenic terpene	Guenther et al. (1995)	Monthly-mean climatology ($2° \times 2.5°$)
di-methyl sulfide (DMS) and methanesulfonic acid (MSA)	Lana et al. (2011)	Monthly-mean climatology ($1° \times 1°$)
Biomass burning[d] SO_2, SO_4, POM[d], and BC	scaled RETROv2 (Duncan et al. 2003) scaled GFEDv3.1 (Randerson et al. 2006) QFED 2.4-r6	Monthly-mean varying (1980 – 1996; $0.3125° \times 0.25°$) Monthly-mean varying (1997 – 2010; $0.3125° \times 0.25°$) Daily (2010 – onwards; $0.3125° \times 0.25°$)
Anthropogenic SO_2	EDGARv4.2 (Energy + Non-Energy) (European Comission 2011)	Annually-varying (1980 – 2008; $0.1° \times 0.1°$)
Anthropogenic SO_4, BC, & POM[d]	AeroCom Phase II (HCA0 v1; Diehl et al. 2012)	Annually-varying (1980 – 2006; $1° \times 1°$)
International Ships (SO_2)	EDGARv4.1 (European Commission 2010)	Annually-varying (1980 – 2005; $1° \times 1°$)
International Ships (SO_4, POM[d], BC)	AeroCom Phase II (HCA0 v1; Diehl et al. 2012) and Eyring et al. (2005)	Annually-varying (1980 – 2007; $1° \times 1°$)
Aircraft (SO_2)	AeroCom Phase II (HCA0 v1; Diehl et al. 2012)	Monthly-varying (1980 – 2006; $1° \times 1.25° \times 72$-levels)

[a]Precursor gasses include SO_2, DMS, and MSA for sulfate aerosol and biogenic terpene for particulate organic matter (POM).

[b]Native resolution of emission inventories in latitude $\times$ longitude in degrees.

[c]Final year of emissions is persisted; however, only degassing volcanic emissions are persisted after 2010.

[d]POM = Particulate Organic Matter = $1.4 \times OC$ as in Textor et al. (2006)

TABLE 2. MERRA-2 Aerosol Optical Depth (AOD) Observing System

Sensor	Temporal Coverage	Description
AVHRR NNR[a]	1980 – August 2002	PATMOS-x radiances over ocean only (PM orbit)
AERONET[b,c]	Station dependent (1999 – October 2014)	AOD from land station network
MISR[c]	February 2000 – June 2014	AOD over bright land surfaces only (albedo > 0.15)
MODIS Terra NNR[a]	March 2000 – onwards (NRT[d])	Collection 5 "Dark Target" land and ocean radiances (AM orbit)
MODIS Aqua NNR[a]	August 2002 – onwards (NRT[d])	Collection 5 "Dark Target" land and ocean radiances (PM orbit)

[a]NNR refers to Neural Net Retrieval algorithm that computes AERONET-calibrated AOD from radiances.

[b]AERONET AOD is converted to 550 nm using an Angström-based interpolation and AOD reported at adjacent channels (500 and 675, typically).

[c]We do not bias-correct AERONET or MISR AOD.

[d]MODIS data is available in Near Real Time (NRT)

TABLE 3. Hourly AERONET Station Statistics* for MERRA-2 and M2REPLAY for 2000 – 2014 (1993 – 1999 in parenthesis).

Station Name	Dominant Aerosol Type	Geographic Region	Latitude	Longitude	Number of co-locations	MERRA-2		M2REPLAY	
						r	RMSE	r	RMSE
Jabiru	smoke	Northern Australia	-12.661	132.893	21917	0.82	0.44	0.68	0.46
Lake Argyle	smoke	Northern Australia	-16.108	128.749	27894	0.81	0.71	0.80	0.48
Alta Floresta	smoke	Amazon Basin	-9.871	-56.104	16942 (1650)	0.92 (0.77)	0.42 (0.72)	0.77	0.62
Bratts Lake	smoke	South-central Canada	50.28	-104.7	16825 (919)	0.85 (0.85)	0.51 (0.46)	0.68	0.52
Rimrock	smoke	Northwestern US	46.487	-116.992	15771	0.75	0.51	0.59	0.56
BSRN BAO Boulder	smoke, polluted	Great Plains US	40.045	-105.006	19545	0.82	0.44	0.70	0.47
Mongu	smoke	South-central Africa	-15.254	23.151	19387 (4831)	0.93 (0.71)	0.32 (0.60)	0.81	0.55
Skukuza	smoke	Southeast Africa	-24.992	31.587	17149 (1803)	0.89 (0.84)	0.34 (0.36)	0.75	0.45
Ilorin	smoke, dust	Sahel	8.32	4.34	15081 (2255)	0.92 (0.76)	0.27 (0.46)	0.71	0.51
Railroad Valley	smoke, dust	Southwest US	38.504	-115.962	23984	0.81	0.60	0.70	0.47
Sevilleta	dust	Southwest US	34.355	-106.885	22144 (8222)	0.84 (0.68)	0.49 (0.53)	0.72	0.46
Arica	dust	Northern Chile	-18.472	-70.313	15906 (3446)	0.67 (0.47)	0.34 (0.48)	0.18	0.82
Saada	dust	Northwest African coast	31.626	-8.156	21290	0.84	0.43	0.82	0.58
Capo Verde	dust	Northwest African coast	16.733	-22.935	24324 (4445)	0.94 (0.84)	0.27 (0.42)	0.81	0.46
Dakar	dust	Sahel	14.394	-16.959	10158	0.92	0.26	0.74	0.49
Banizoumbou	dust	Sahel	13.541	2.665	35365 (3958)	0.80 (0.64)	0.46 (0.66)	0.65	0.58
IER Cinzana	dust	Sahel	13.278	-5.934	28947	0.82	0.40	0.68	0.51
Dalanzadgad	dust	Inner Mongolia	43.577	104.419	18701 (2692)	0.82 (0.75)	0.56 (0.64)	0.70	0.60
Solar Village	dust	Saudi Arabia	24.907	46.397	33837 (2704)	0.94 (0.92)	0.33 (0.28)	0.73	0.51
Sede Boker	dust	Israel	30.855	34.782	39027 (3225)	0.89 (0.80)	0.45 (0.42)	0.73	0.48
IMS-METU-ERDEMLI	dust	Turkey	36.565	34.255	16444	0.83	0.32	0.67	0.55
Nes Ziona	dust	Israel	31.922	34.789	23640	0.89	0.31	0.70	0.43
GSFC	polluted	Northeastern US	38.992	-76.84	24331 (6479)	0.94 (0.84)	0.50 (0.51)	0.79	0.52
MD Science Center	polluted	Northeastern US	39.283	-76.617	19285 (528)	0.89 (0.86)	0.52 (0.56)	0.72	0.56
Wallops	polluted	Northeastern US	37.942	-75.475	10652 (1924)	0.93 (0.90)	0.48 (0.46)	0.79	0.50
Mexico City	polluted	Central Mexico	19.334	-99.182	15830 (1045)	0.58 (0.54)	0.76 (0.86)	0.54	0.97

*Statistics from natural log-transformed hourly AOD co-locations: r is Pearson's correlation, RMSE is root mean square of the differences.

TABLE 4. **Table 3 Continued**: Hourly AERONET Station Statistics* for MERRA-2 and M2REPLAY for 2000 – 2014 (1993 – 1999 in parenthesis).

Station Name	Dominant Aerosol Type	Geographic Region	Latitude	Longitude	Number of co-locations	MERRA-2 r	MERRA-2 RMSE	M2REPLAY r	M2REPLAY RMSE
Lille	polluted	France	50.612	3.142	7421	0.85	0.34	0.59	0.46
Avignon	polluted	France	43.933	4.878	23913 (57)	0.81 (0.33)	0.39 (0.56)	0.72	0.44
Carpentras	polluted	France	44.083	5.058	21486	0.85	0.42	0.72	0.47
Ispra	polluted	Italy	45.803	8.627	13896	0.85	0.46	0.69	0.68
Rome Tor Vergata	polluted	Italy	41.84	12.647	2972	0.71	0.39	0.69	0.46
Venise	polluted	Italy	45.314	12.508	10495 (873)	0.84 (0.82)	0.51 (0.48)	0.63	0.59
Moldova	polluted	Black Sea coast	47	28.816	20092 (269)	0.87 (0.73)	0.35 (0.43)	0.67	0.45
Evora	polluted, dust	Spain	38.568	-7.912	10771	0.87	0.45	0.78	0.47
CEILAP-BA	polluted, dust	Argentina	-34.567	-58.5	18102 (454)	0.73 (0.73)	0.52 (0.28)	0.50	0.51
Kanpur	polluted, dust	Central India	26.513	80.232	22029	0.85	0.28	0.43	0.80
XiangHe	polluted, dust	Northeastern China	39.977	116.381	17730	0.88	0.47	0.79	0.70
Beijing	polluted, dust	Northeastern China	39.754	116.962	22041	0.84	0.56	0.78	0.73
La Parguera	clean marine, dust	Caribbean	17.97	-67.045	19138	0.90	0.34	0.73	0.46
Cart Site	clean continental	Great Plains US	36.607	-97.486	20312 (2910)	0.89 (0.85)	0.39 (0.39)	0.79	0.39
Bondville	clean continental	Great Plains US	40.053	-88.372	15356 (3397)	0.89 (0.78)	0.45 (0.50)	0.73	0.49
Ascension Island	clean marine	South Atlantic	-7.976	-14.415	10794 (992)	0.87 (0.49)	0.27 (0.56)	0.55	0.49

*Statistics from natural log-transformed hourly AOD co-locations: r is Pearson's correlation, RMSE is root mean square of the differences.

TABLE 5. AOD and AAOD Comparisons Between Reanalyses, Models, and Observations

	MERRA-2[a,g]	MERRAero[a,g]	NAAPS[b,h]	MACC[c,i]	AeroCom Phase I[d,g]	Yu et al. Models[e]	Yu et al. Obs.[f]
Sulfate	0.040 ± 0.004	0.039 ± 0.004	–	–	0.034 ± 0.011	–	–
Black Carbon	0.006 ± 0.001	0.006 ± 0.001	–	–	0.004 ± 0.002	–	–
Organic Carbon	0.022 ± 0.006	0.025 ± 0.007	–	–	0.019 ± 0.007	–	–
Dust	0.030 ± 0.010	0.026 ± 0.009	0.039	0.043 ± 0.014	0.032 ± 0.014	–	–
Sea Salt	0.041 ± 0.002	0.034 ± 0.002	0.035	0.055 ± 0.016	0.030 ± 0.015	–	–
Total AOD	**0.140** ± 0.013	**0.130** ± 0.015	**0.137**	**0.180** ± 0.030	**0.127** ± 0.025	**0.129** ± 0.033	**0.162** ± 0.023
Fine Mode[g,h,i]	0.068 ± 0.008	0.070 ± 0.011	0.064	0.082	0.063 ± 0.016	–	–
Coarse Mode[j]	0.072 ± 0.010	0.060 ± 0.009	0.073	0.098	0.061 ± 0.024	–	–
AAOD [k]	**0.007** ± 0.001	**0.009** ± 0.001	–	**0.008** ± 0.002	**0.005** ± 0.002	–	–

[a]Climatological global area-weighted average $\pm$ standard deviation of monthly AOD for Y2003 – Y2010.

[b]NAAPS aerosol reanalysis calculated for Y2003 – Y2010 from annual average AOD (Lynch et al. 2016).

[c]MACC Y2003 – Y2010 global mean and uncertainty from Bellouin et al. (2013).

[d]AeroCom Phase I multi-model median and standard deviation from Kinne et al. (2006).

[e]Median and standard deviation from 4 global models considered in Yu et al. (2006).

[f]Median and standard deviation from MODIS Terra, MISR, and combinations of these satellite datasets with GOCART from Yu et al. (2006).

[g]Fine mode AOD is the sum of sulfate, organic carbon, and black carbon AOD for MERRA-2, MERRAero, and AeroCom.

[h]NAAPS fine mode AOD is the sum of anthropogenic fine plus biomass burning AOD.

[i]MACC fine mode AOD is the sum of anthropogenic plus fine mode natural AOD.

[j]Coarse mode AOD is the sum of dust and sea salt.

[k]Bellouin et al. (2013) use an algorithm to determine aerosol speciation and AAOD, not the native MACC AAOD.
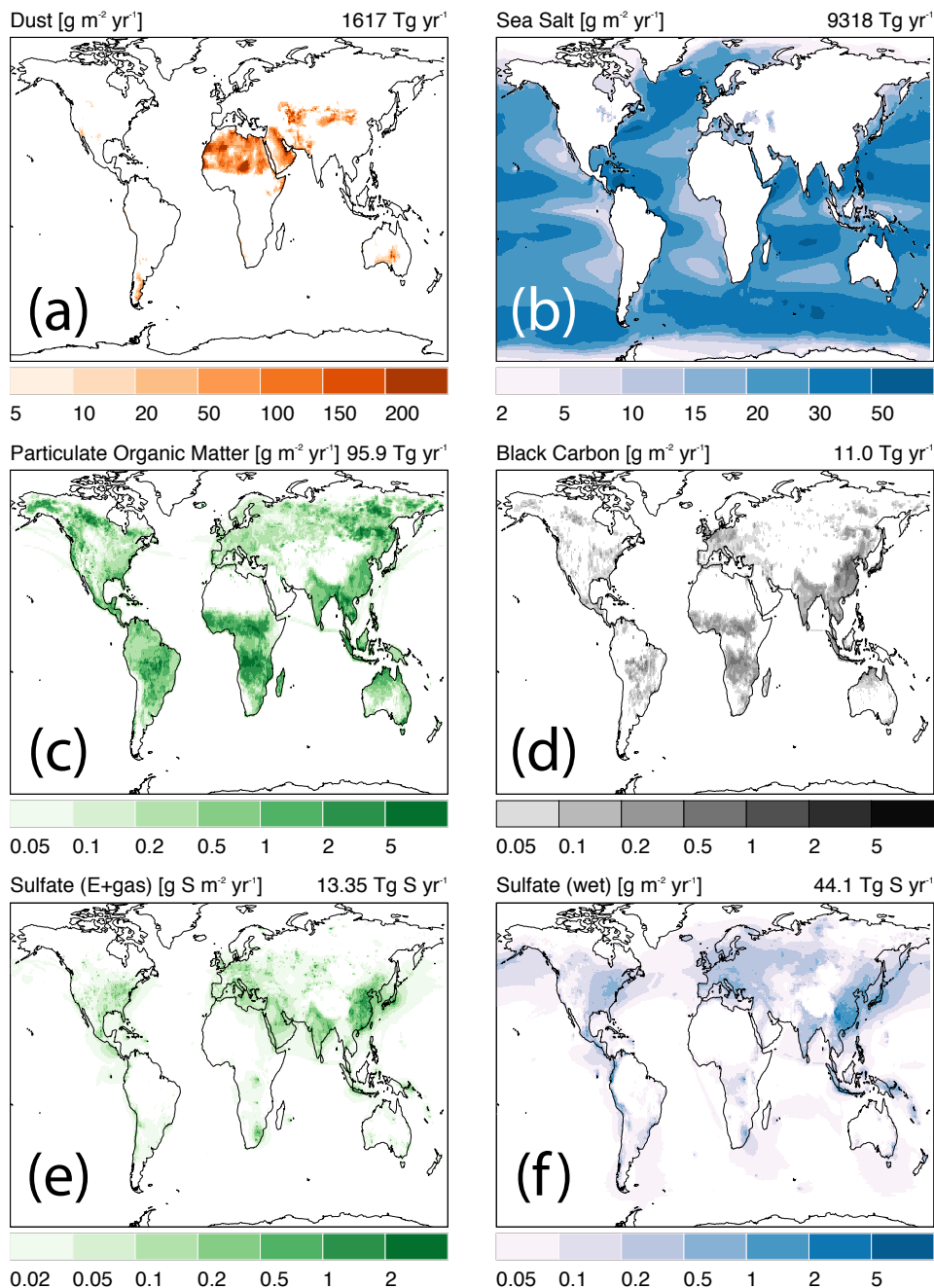
FIG. 1. Annual-mean aerosol emissions climatology (2000 – 2014) from MERRA-2. Emissions are shown for (a) dust, (b) sea salt, (c) particulate organic matter (POM), (d) black carbon (BC), (e) primary sulfate (SO_4) and sulfate from oxidation of sulfur dioxide gas (SO_2), and (f) SO_4 from aqueous production. We define POM = 1.4 × OC, where OC is organic carbon (Textor et al. 2006). Emissions are from all sectors (fossil fuel, biofuel, biomass burning, and biogenic, if applicable). The global, annual mean emissions are given on each panel.
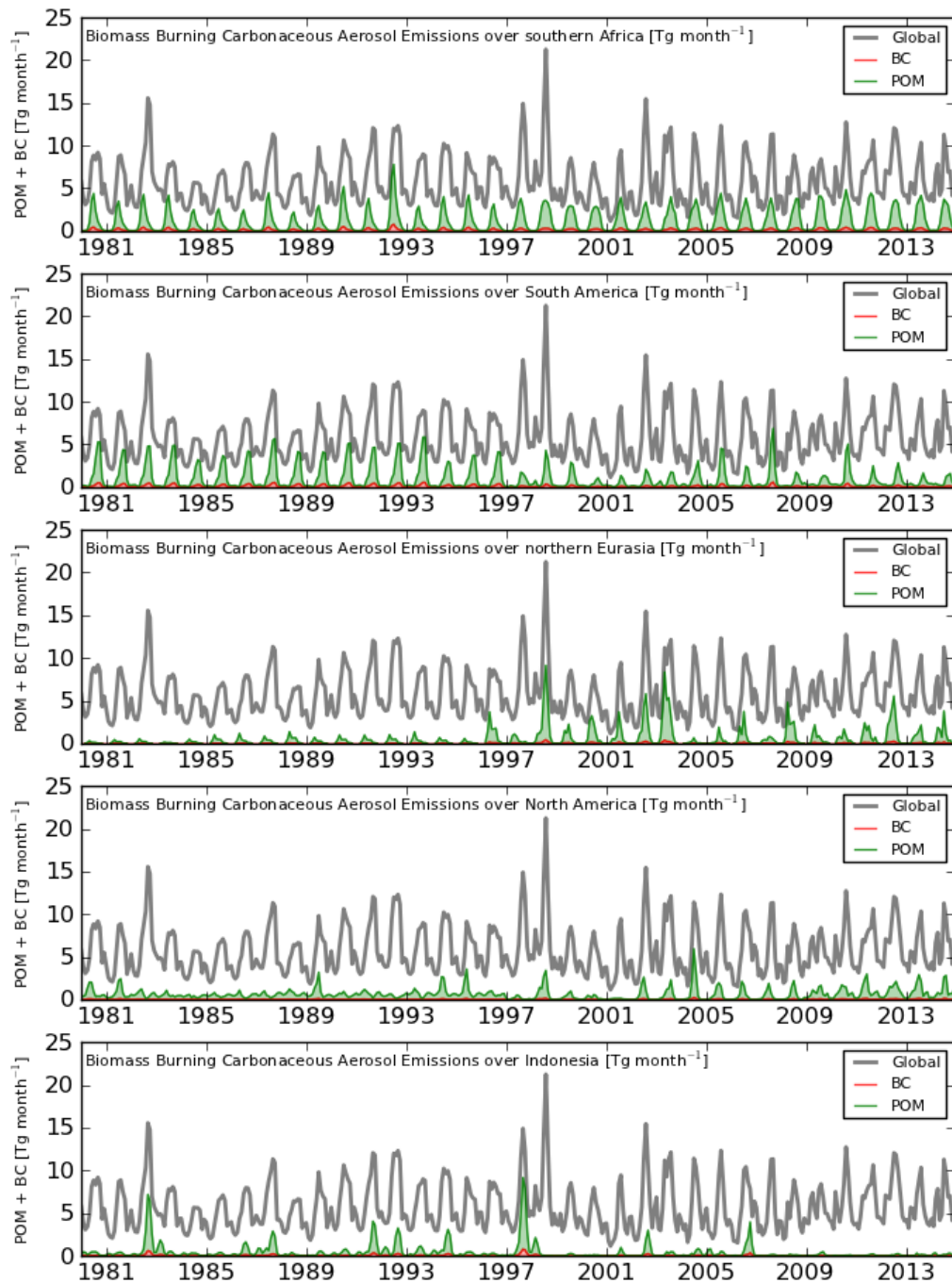
FIG. 2. Timeseries of carbonaceous aerosol emissions (BC + POM) in Tg month^{-1} from biomass burning sources averaged globally (grey line; same in all panels) and over several major source regions (panels). For the regional emissions, red and green shading indicates the relative contribution of BC and POM, respectively, to the total carbonaceous aerosol emissions; the total carbonaceous aerosol emissions are the sum of the POM and BC contributions (i.e. stacked area plots are shown, but POM emissions far exceed BC emissions). Comparing the regional and global timeseries, we show regional emissions events of global importance, such as the 1997 Indonesian fires.
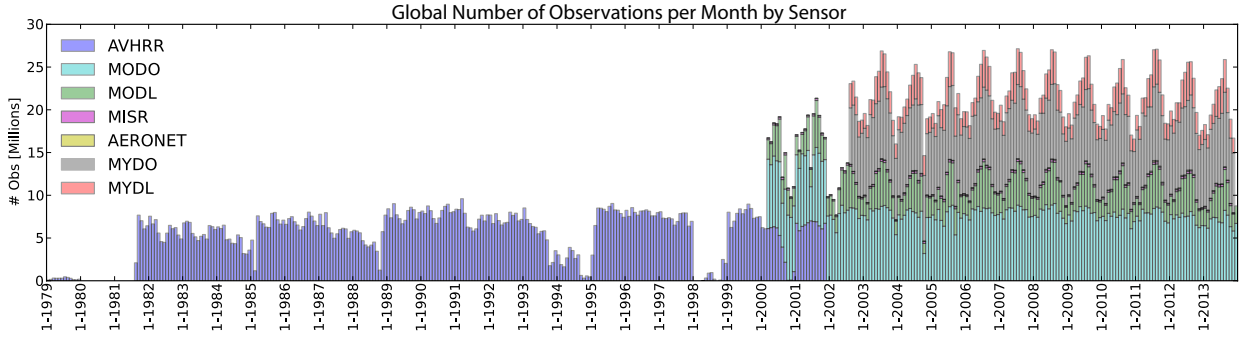
FIG. 3. Total global monthly number of observations from AVHRR NNR, MODIS Terra over land (MODL NNR) and ocean (MODO NNR), MODIS Aqua over land (MYDL NNR) and ocean (MYDO NNR), MISR over bright surfaces (deserts), and AERONET, where NNR is the bias-corrected neural net retrieved AOD. Note the following: (1) AVHRR observations are only over the ocean. (2) Stronger cloud contamination in the Southern Hemisphere relative to the Northern Hemisphere imparts a seasonal variation on the data volume for satellite sensors during the EOS-period. (3) Data counts from MISR and AERONET are more clearly seen in the Supplementary Figures.
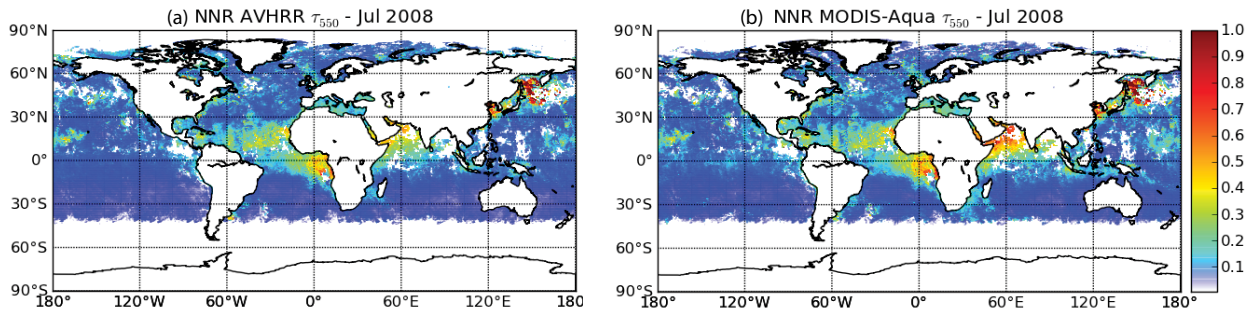
FIG. 4. (a) AVHRR NNR AOD in July, 2008 where there is overlap with (b) MODIS Aqua NNR AOD to show the consistency applied to the bias-corrected AOD by the NNR algorithm.
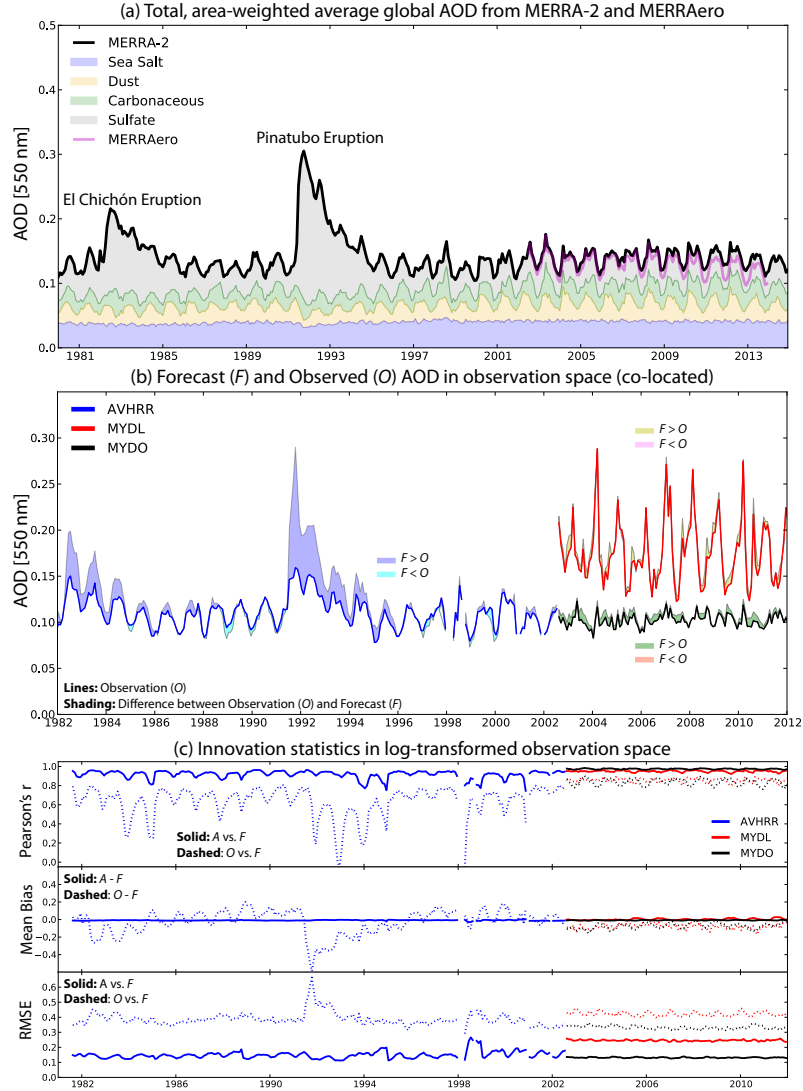
FIG. 5. (a) Global, area-weighted average monthly-mean 550 nm AOD from MERRA-2 (black line) and MERRAero (magenta) and contributions from various aerosol species (shading, MERRA-2 only). Carbonaceous aerosol AOD (green) is the sum of organic plus black carbon AOD. Note that only the total AOD (thick black line) is directly constrained by the aerosol assimilation; aerosol speciation depends strongly on emission and loss processes in the model. (b) Monthly-mean, co-located in time and space (pair-wise in observation space) comparison of the 3-hour model forecast (F) and observations (O) for AVHRR NNR, MODIS Aqua land NNR (MYDL), and MODIS Aqua ocean NNR (MYDO). Here the lines represent the observed AOD, and shading indicates the difference between F and O. (c) Monthly-mean statistics computed in log-transformed observation space. Statistics comparing O and F are shown as dashed lines, and relationships between the observations and analysis (A) are shown as solid lines for the same sensors as in (b). Supplementary figures show the statistics for the other sensors in the MERRA-2 aerosol observing system (Table 2). The log-space errors (standard deviation) in AOD are related to linear errors according to Eq. 7. Note that we expect the monthly mean statistics presented here to outperform statistics sampled at higher time frequencies.
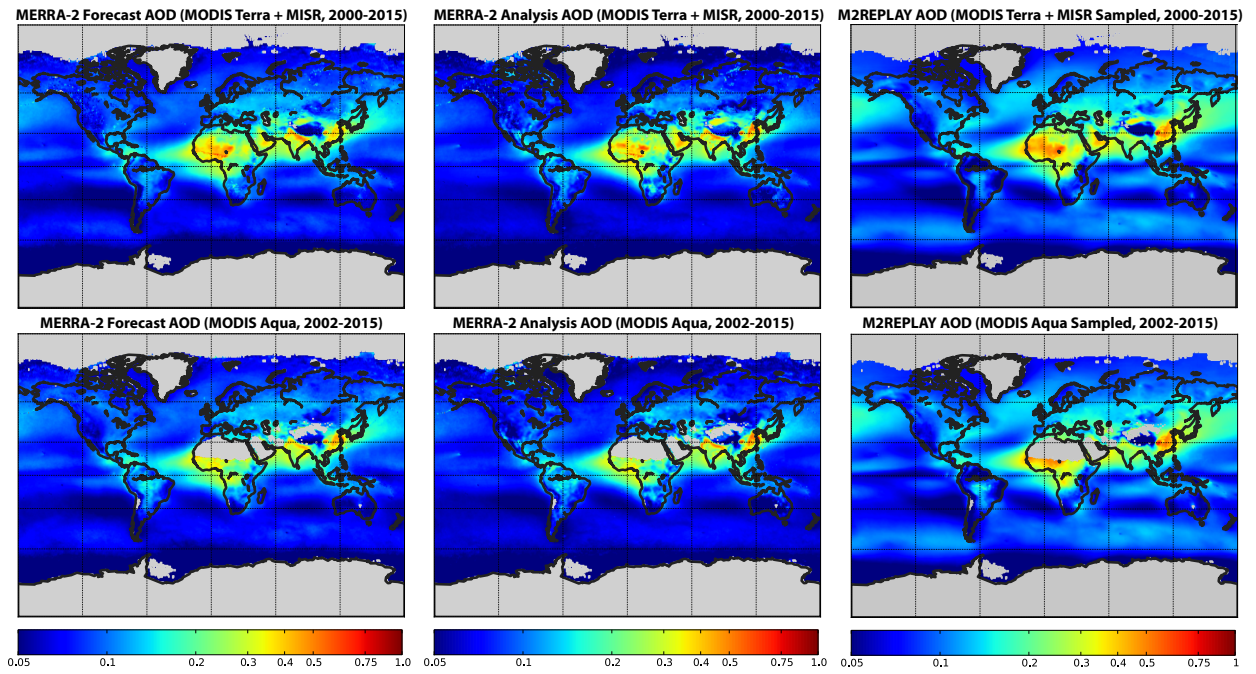
FIG. 6. Climatological annual-mean 3-hr forecast (left column), analysis (center column), and control run (M2REPLAY; right column) AOD for MODIS Terra plus MISR over bright surfaces (top row) and MODIS Aqua (bottom row). Grey regions indicate no data, and the control run has been sampled with the day-time orbit of the relevant sensors. Differences between the left and center columns indicate where the the analysis differs from the forecast; in most places, the analysis AOD strongly resembles the forecast. Differences between the center and right column indicate the overall impact of the analysis on simulated AOD since we compare to a control simulation without AOD assimilation.
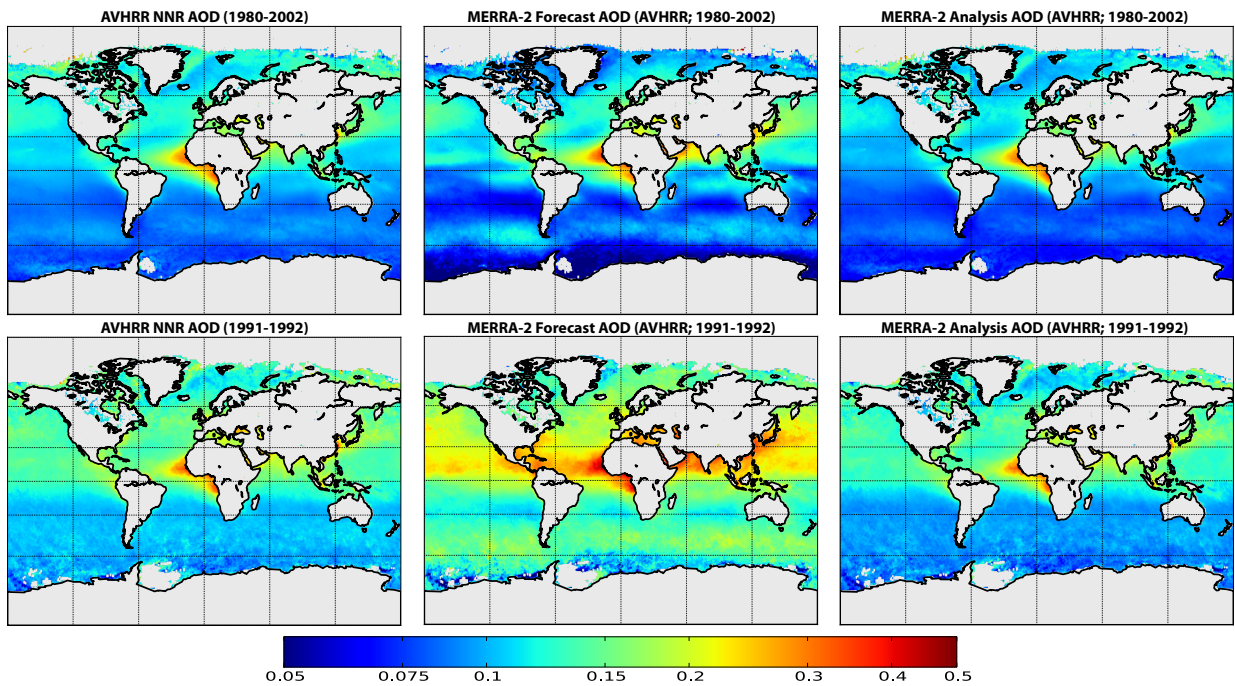
FIG. 7. Observed (*O*; left column), forecasted (*F*; center column) and analyzed (*A*; right column) AOD in AVHRR observation-space for (top row) 1980 – 2002 and (bottom row) 1991–1992 after the Pinatubo eruption. Grey regions indicate regions with no data. Note that the bias-corrected observations in the left column are from the AVHRR NNR. Forecasted AOD after Pinatubo (bottom row, center column) is generally higher than the observations and analyzed AOD fields, implying negative AOD increments during this time period. We also note possible cloud contamination in the high latitude observations from AVHRR which may impact the analysis.
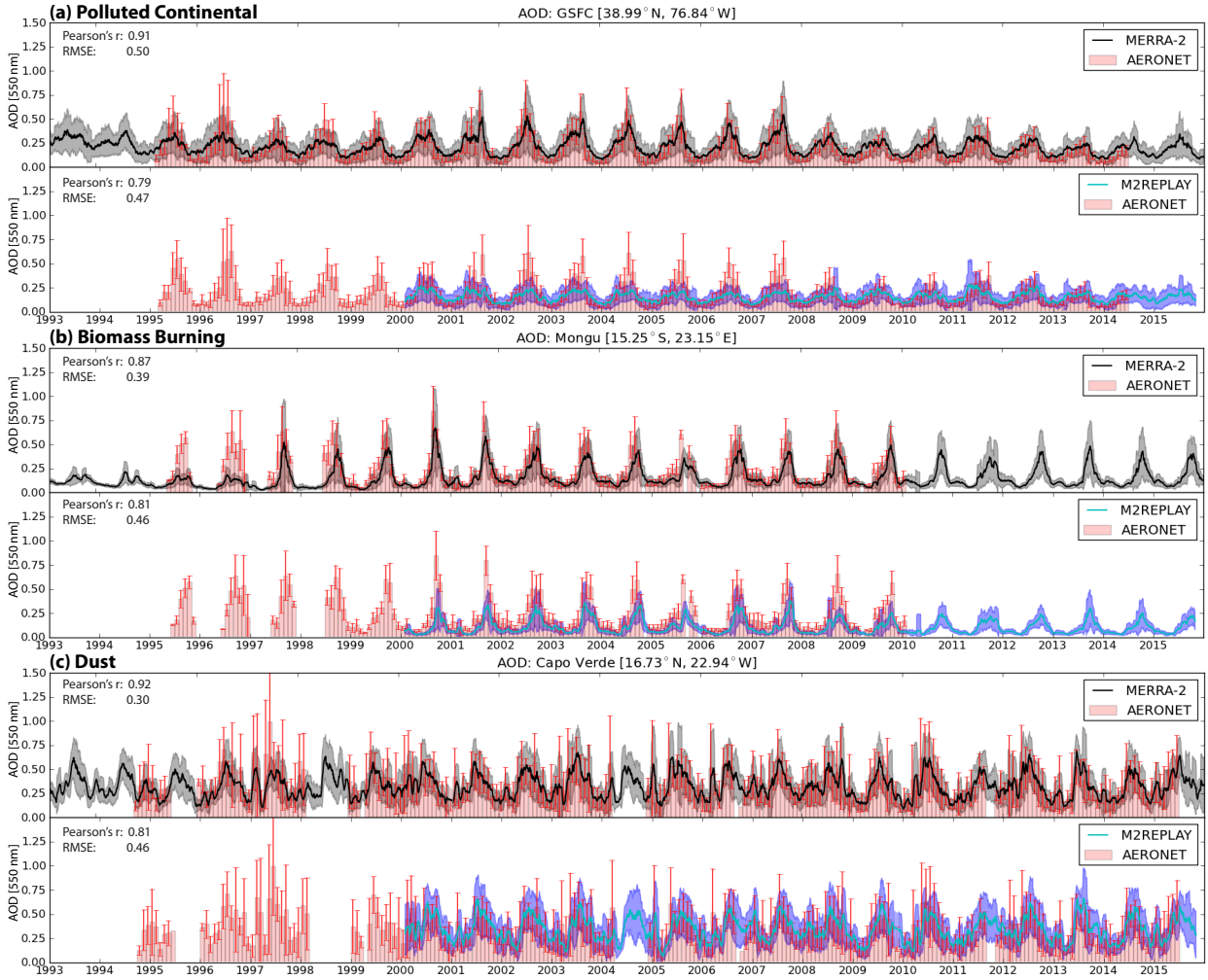
FIG. 8. Timeseries of AOD at three AERONET sites representing (a) polluted continental (GSFC, US East Coast), (b) biomass burning (Mongu, central Africa), and (c) dust (Capo Verde, northwest African coast) dominated regions. Red bars are AERONET monthly mean AOD with error bars indicating the standard deviation of the daily observed AOD. Black and blue lines are the rolling mean AOD from MERRA-2 and the control run (M2REPLAY), respectively, with shading representing the rolling standard deviation. Correlation coefficients (*r*) are based on co-located hourly data for the entire timeseries shown in each panel.
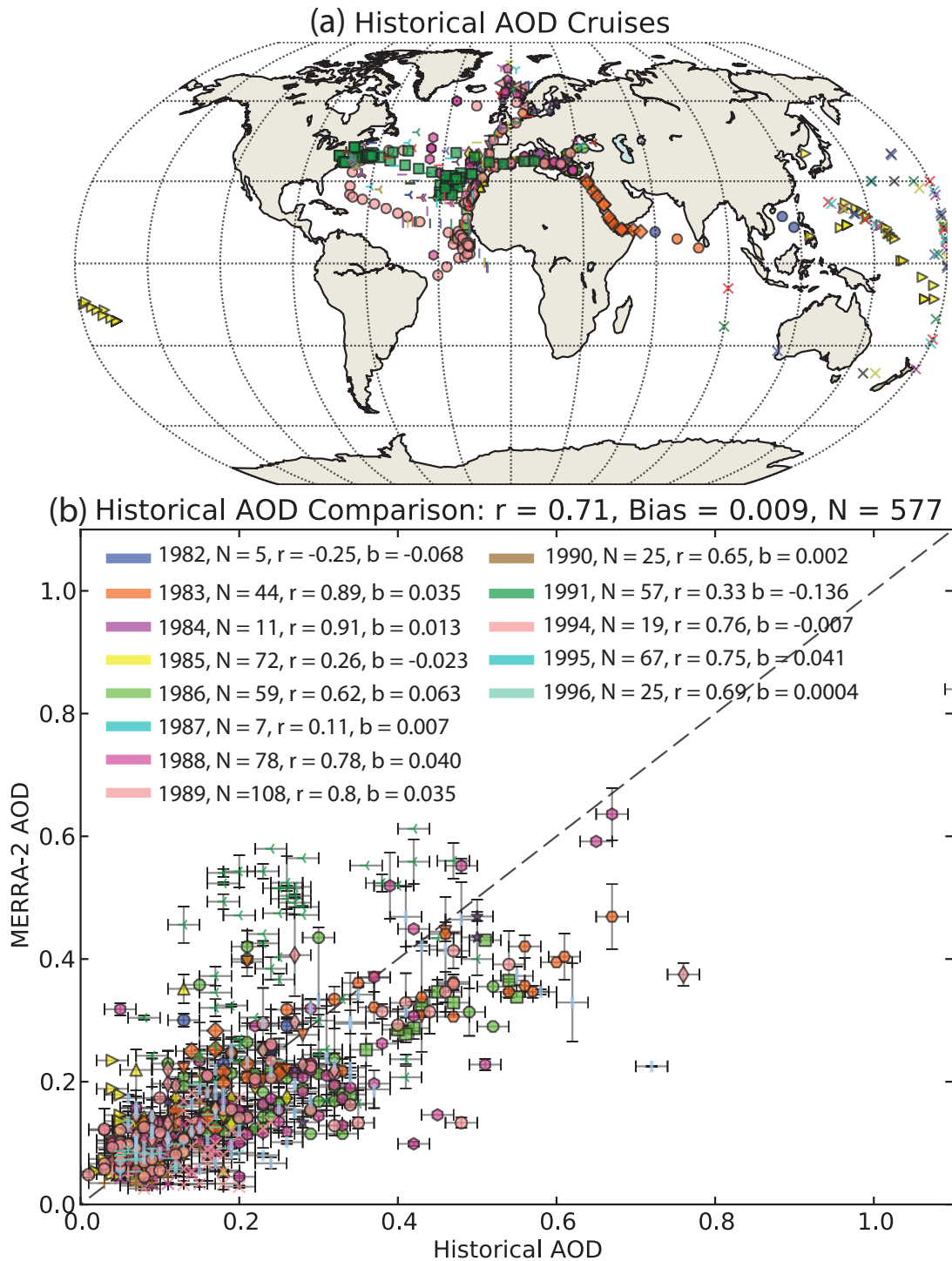
FIG. 9. Comparison of MERRA-2 and historical shipborne AOD observations. (a) Map showing the location of the ship cruises spanning the period 1982–1996. (b) Scatter plot of AOD spanning various years (indicated by different colors) and different cruises (indicated by marker shape). X-axis error bars represent the assumed observed AOD error of 0.02, and y-axis error bars are the standard deviation of the MERRA-2 AOD used to make the morning or afternoon averages (see text for sampling strategy).
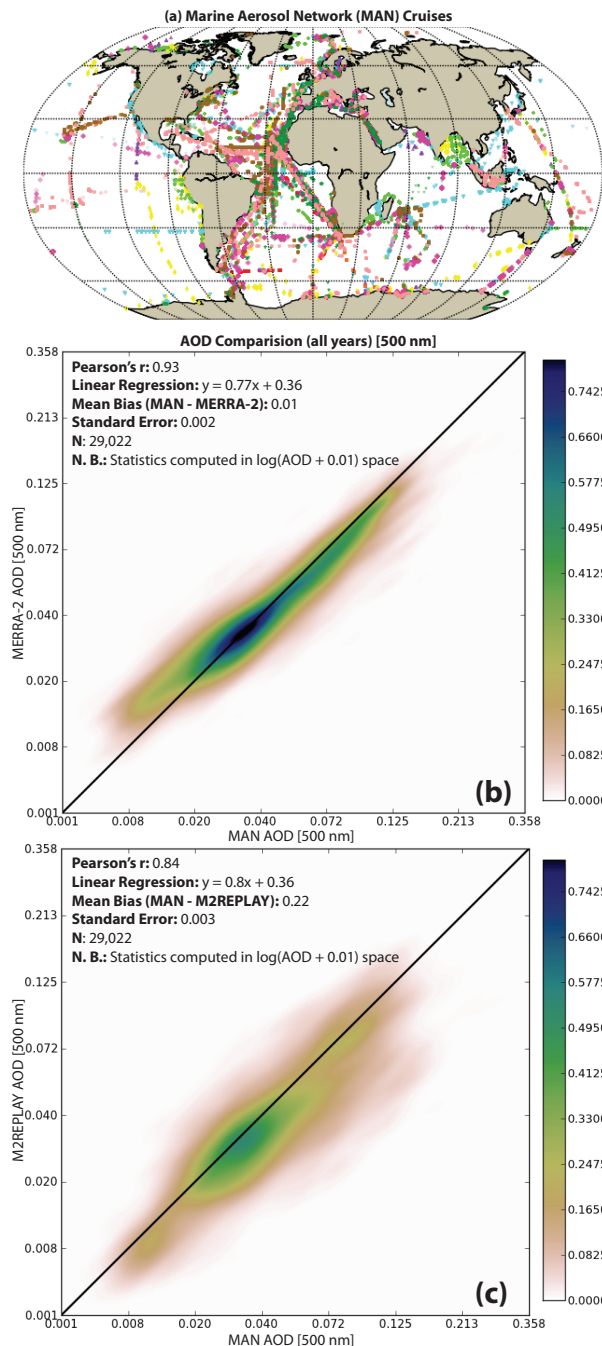
FIG. 10. (a) Maritime Aerosol Network (MAN) cruises 2004 – 2015, color coded by year (see the MAN website for information on specific cruises). (b) Joint PDF comparison of co-located MAN-observed and MERRA-2 AOD for the same period. (c) Same as (b) but for the M2REPLAY control simulation without AOD assimilation. Note that statistics are calculated in natural log-transformed AOD space, the colorbar represents probability density, and the x- and y-axes have been re-labeled in linear AOD space for clarity.
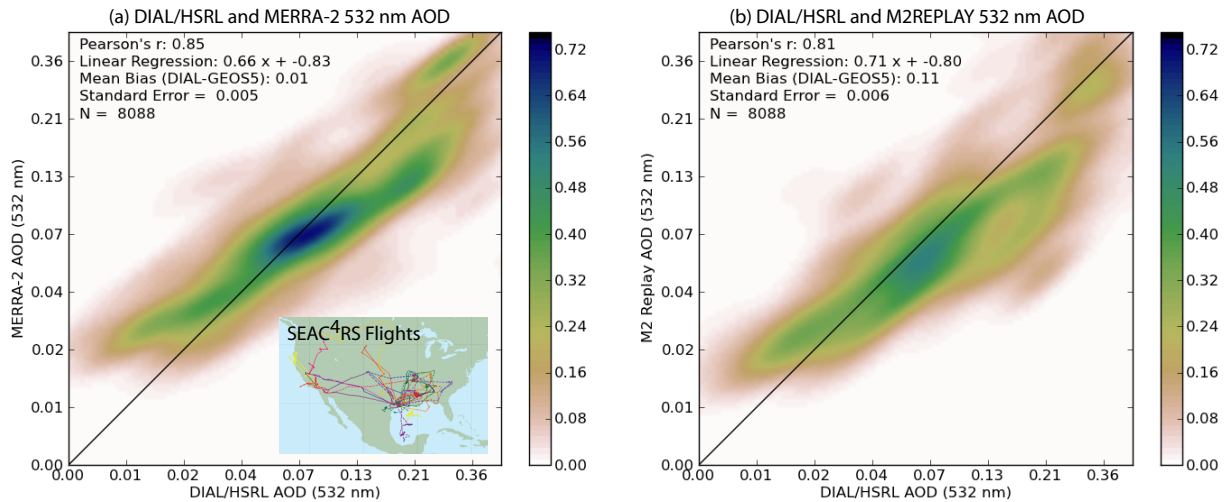
(a) DIAL/HSRL and MERRA-2 532 nm AOD

Pearson's r: 0.85
Linear Regression: 0.66 x + -0.83
Mean Bias (DIAL-GEOS5): 0.01
Standard Error = 0.005
N = 8088

(b) DIAL/HSRL and M2REPLAY 532 nm AOD

Pearson's r: 0.81
Linear Regression: 0.71 x + -0.80
Mean Bias (DIAL-GEOS5): 0.11
Standard Error = 0.006
N = 8088

FIG. 11. Comparison of AOD observed during the NASA SEAC^{4}RS campaign (August – September, 2013) over the southeastern and western continental United States (see inset map). We compare 532 nm AOD observed by the DIAL/HSRL instrument aboard the NASA DC8 aircraft for the entire campaign to (a) MERRA-2 and (b) M2REPLAY sampled along the flight paths and calculated over the same portion of the column as reported by the instrument. Statistics are reported in natural log-transformed AOD space, the colorbar represents probability density, and the x- and y-axes have been re-labeled in linear AOD space for clarity.
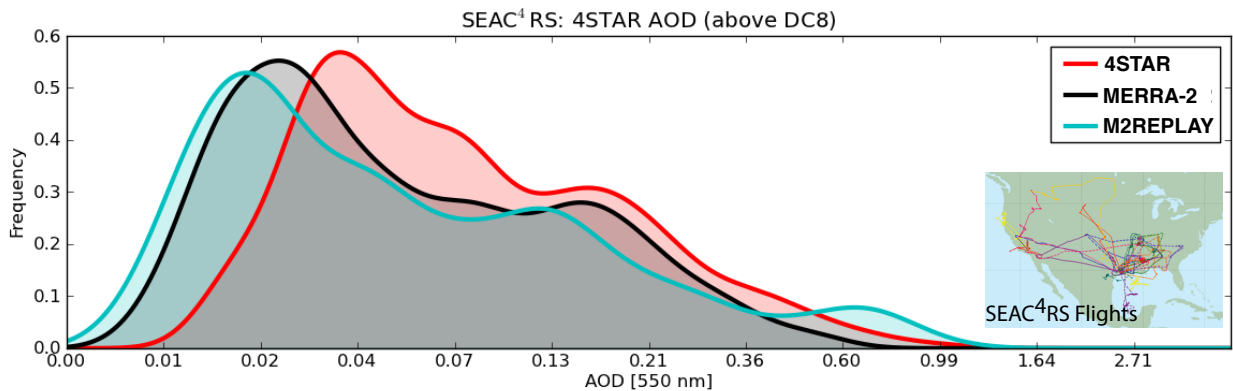

FIG. 12. PDFs of AOD from the NASA SEAC^{4}RS campaign (August – September, 2013) over the southeastern and western continental United States (see inset map). We compare 550 nm AOD PDFs observed by the 4STAR instrument (red) aboard the NASA DC8 aircraft for the entire campaign to MERRA-2 (black) and M2REPLAY (blue) sampled along the flight paths and calculated over the same portion of the column (only above the aircraft) as reported by the instrument. Note the logarithmic spacing of the x-axis, and that the differences in modeled and observed AOD for the lowest maximum is within the reported instrumental error ($\pm$0.02 for SEAC^{4}RS).
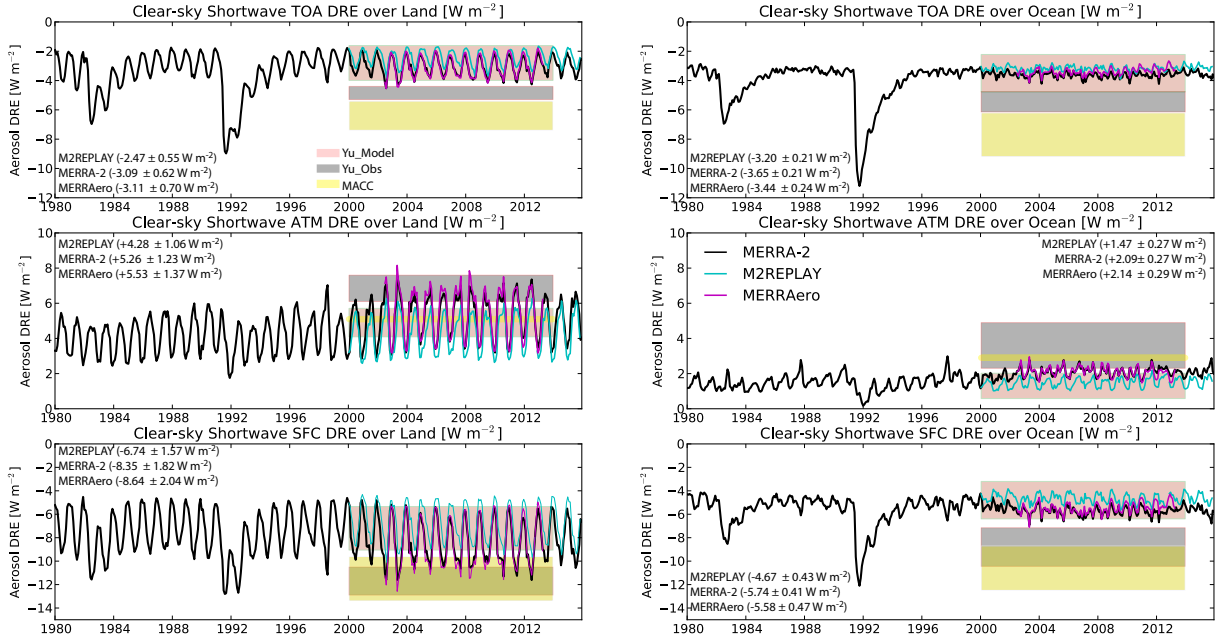
FIG. 13. Timeseries of the global monthly mean clear-sky shortwave aerosol Direct Radiative Effect (DRE) in W m^{-2} over land (left column) and ocean (right column). DRE is shown for (top row) top-of-the atmosphere (TOA), (middle row) atmosphere (ATM), and (bottom row) surface (SFC) DRE where TOA = SFC + ATM. Full time series of DRE are shown for MERRA-2 (black line), M2REPLAY (cyan line), and MERRAero (magenta line). Red shading is the Yu et al. (2006) multi-model range (median $\pm$ standard deviation), and grey shading is the satellite observation-derived DRE range (median $\pm$ standard deviation) from the same study. Yellow shading is the estimated DRE (2003-2010 mean $\pm$ uncertainty) derived from the MACC aerosol-reanalysis AOD (Bellouin et al. 2013). Note that the ATM forcing in Bellouin et al. (2013) is reported without an uncertainty range, so the reported mean value is shown as a yellow line. The 2003 – 2010 mean and standard deviation DRE for MERRA-2, M2REPLAY, and MERRAero are given on each panel. See Supplementary Table 3 for ocean- and land-averaged AOD, AAOD, and DRE efficiency.